Rain or Shine? Optimal Utility Pricing under Different Weather Patterns

By GUOZHEN (GORDON) JI

THE LATEST VERSION OF THIS PAPER CAN BE FOUND HERE

As climate change amplifies more volatile weather patterns, water utilities face increasing difficulty in simultaneously ensuring revenue feasibility, promoting water conservation, and protecting low-income consumers. This paper tests and concludes that price alone cannot achieve these competing policy goals under different weather patterns. Using granular household data from Austin, TX, and a structural demand model enhanced with satellite imagery-derived vegetation index, I find that because high-water users exist across all income levels, traditional tiered pricing doesn't work as intended. Furthermore, higher-income households—who are both weather-sensitive and surprisingly price-elastic—complicate the utility's ability to achieve its distributional objectives while meeting the conservation target. When high-demand conditions (e.g., drought) make conservation measures necessary, low-income families experience an average welfare loss of \$74 per month. This highlights the necessity of complementary policies to achieve distributional goals when demand increases. For example, a program encouraging households to convert 30% of their lawns to water-saving landscapes (zeroscaping/xeriscaping) could generate approximately \$70 per month in welfare for the lowest-income families, nearly offsetting the financial burden imposed by conservation policies during droughts.[*] *Keywords: Ramsey Pricing, Multi-part Tariff, Demand Volatility, Distributional Effect*

I. Introduction

The increasing variability of weather patterns, exacerbated by climate change, poses a significant challenge to the pricing of essential utility services, particularly water. Urban water utilities face a complex balancing act, primarily managed through their rate structures: they must ensure financial sustainability through cost recovery, promote resource conservation amid dwindling supplies, and maintain equitable access for all consumers,

[*] I am deeply grateful to my advisor, Dr. Eugenio Miravete, whose inspiration and guidance were essential throughout this project. His encouragement to pursue this topic and acquire the necessary data, along with his invaluable feedback over the years, was instrumental to this work. I am also grateful to my committee members—Dr. Robert Town, Dr. Jackson Dorsey, and Dr. Andrey Ordin—for their insightful comments on the research and helpful feedback on its presentation. I extend my sincere thanks to Dr. Daniel Ackerberg, Dr. Jorge Balat, Dr. Victoria Monroe, Dr Sheila Olmstead, and Dr. Jori Barash for taking the time to give this project guidance. I would also like to thank Austin Water for generously providing the data and for their trust in this project. Without their support, this research would not have been possible. Finally, I thank Priyansh Dhandha for his immense contributions as a research assistant on the empirical work.

especially low-income households. This study focuses on these challenges within the context of Austin, Texas—a rapidly growing metropolitan area known for its highly variable precipitation. The primary objective of this paper is to develop and empirically implement an optimal water pricing framework, grounded in Ramsey principles and tailored for an Increasing Block Price (IBP) tariff, that explicitly accounts for and remains resilient while achieving policy goals related to financial feasibility, conservation, and distributional equity under extreme weather patterns.

The theoretical underpinning for many utility pricing schemes stems from Ramsey (1927), who proposed "second-best" solutions to cover large fixed costs while minimizing societal welfare loss, typically through an inverse elasticity rule. In practice, utilities often adopt IBP tariffs, where prices rise with consumption, to encourage conservation among high-volume users while ensuring affordability for basic needs. However, the efficacy of IBP structures is predicated on the assumption that high consumption correlates with high income and low price elasticity—an assumption severely tested by unpredictable weather. A critical operational constraint is that utility prices are typically preset for an entire year, lacking the flexibility to adjust to monthly or seasonal weather variations. This means a single, static IBP structure must bear the burden of achieving all policy goals, from mitigating excessive usage during dry months to offsetting revenue loss during wet months, all while addressing distributional concerns. This paper first offers reduced-form evidence of consumer sensitivity to precipitation and then structurally estimates demand under nonlinear IBP using a Discrete/Continuous Choice (DCC) model.

Previous literature has explored the complex nature of estimating public utility demand and the welfare impacts of IBP optimization. Hewitt and Hanemann (1995) and Olmstead, Hanemann and Stavins (2007) use DCC models to estimate price elasticity in urban water demand. Some studies, such as Castro-Rodríguez, María Da-Rocha and Delicado (2002) and Nataraj and Hanemann (2011), have shown that demand is responsive and that specific IBP changes can generate welfare improvements. In particular, Szabo (2015), examining South Africa's free water policy, demonstrated that nonlinear pricing alone can improve welfare and conservation goals without relying on ad-hoc subsidies. Though these studies have demonstrated that IBP can meet the conservation policy while maximizing consumer and producer surplus, the distributional effect of IBP has been less discussed, likely due to the lack of granular data. Recently, with more detailed data, research has been highlighting the potential shortcomings of relying heavily on IBP as a policy instrument, particularly regarding the distributional goal. Borenstein (2009) and Ito (2014) found that consumers appear to react more to average prices than to the marginal prices of multiple tiers, and households that usually have less understanding of the complex pricing structure are likely to have lower income levels. Ruijs (2009) and Echeverri (2023), while acknowledging that consumers respond to IBP changes, directly evaluate the distributional effect of IBP restructuring and concluded that higher-income households usually gain higher welfare through IBP price changes.

This paper fills a gap in the literature by exploring the effectiveness of IBP as a price instrument under the strain of different weather patterns. I align closely with Wolak

(2016), who employed a Ramsey-style model to determine an optimal IBP structure for welfare maximization. However, to my knowledge, this paper is the first to develop and empirically apply a Ramsey-style optimal pricing model that uniquely incorporates different exogenous weather patterns—both in terms of mean shifts and variance shifts—to "stress test" the efficacy of IBP, and to prepare the utility's policy goals under all types of weather conditions. Furthermore, the demand estimation is enhanced by the novel integration of high-resolution (10m $\times$ 10m) satellite-derived Normalized Difference Vegetation Index (NDVI) data. My descriptive analysis and the structural DCC model reveal that, contrary to common assumptions, high-quantity users exist across all income strata and that higher-income households can exhibit greater price elasticity, particularly when weather pushes the demand curve rightward. This finding introduces significant challenges for the utility relying solely on IBP to achieve its distributional goals when weather pushes the demand curve to the right, while conserving resources.

This research also relates to the industrial organization literature on firms' responses to weather-induced demand volatility, such as Lin, Schmid and Weisbach (2017) and Baumgartner et al. (2022), and to research about the social consequences of real-time electricity pricing by Holland and Mansur (2008) and the heterogeneities of social marginal cost caused by weather by Borenstein and Bushnell (2022). The latter two papers about electricity pricing also discussed the utility's conservation policy goal under volatile demand, but they focused on how frequently pricing should change in response to volatile demand, a unique feature that is more feasible in electricity pricing.[1] This paper takes the inflexibility of price change within a year as an exogenous political constraint, and extends the existing paper by adding the specific concern regarding the distributional effects of pricing, on top of the environmental concerns, by exploiting the weather-induced heterogeneity of urban residential water demand.

Lastly, this paper contributes to a growing literature evaluating the distributional effects of climate-related policies (Parry and Williams III (2010), Goulder et al. (2019), Känzig (2023))in both the field of industrial organization and macroeconomics, which has often found such policies to be regressive. Instead of targeting a specific green policy, this paper extends this strand of literature by evaluating how much of the regressiveness could be controlled by the utility's policy goal, and how much of it is exogenous and caused by weather. This disentanglement allows me to quantify the shadow cost of the conservation policy goal towards lower-income households and the trade-off between the policies of conservation and the distributional effect through price under various weather patterns.

I find that when weather pushes the demand curve to the right (e.g., drought), price alone struggles to improve, or even maintain, the distributional effect due to the binding conservation constraint. When demand shifts rightward due to drought or high weather variance, the conservation constraint imposes an average welfare loss of \$74.2 and \$74.9 per household per month for the lowest-income stratum, respectively. However, when demand shifts leftward due to more precipitation or low weather variance, the optimal

[1] Due to existing industry norms (like peak and valley pricing), consumers are more likely to accept real-time electricity pricing, and the utility will face less political consequences, as oppose to water utility.

price helps the lower-income stratum gain an average welfare of \$45.3 and \$48.5, respectively.[2] The estimated shadow cost of the conservation policy during drought can inform the design of complementary policies to achieve distributional objectives. For example, a 30% household zeroscaping/xeriscaping[3] effort could generate \$75 in welfare per household per month for the lowest-income stratum, almost nullifying the welfare loss from the conservation policy's shadow cost.

The remainder of this paper is structured as follows. Section II provides background on water utility pricing in Austin, TX, discusses the inherent challenges posed by volatile weather, describes the datasets utilized, and presents descriptive evidence. Section III details the development of the structural demand model, a Discrete/Continuous Choice model, explains the estimation strategy incorporating household characteristics, weather variables, and NDVI data, and presents the key estimation results, including price and income elasticities, their heterogeneities, and how weather shifts the demand curve. Section IV establishes the theoretical Ramsey pricing model, extends it to incorporate conservation policy constraints, and outlines the empirical model used to determine optimal IBP parameters. Section V presents the counterfactual analysis, where I estimate optimal prices and their welfare consequences under various shifts in precipitation patterns (both mean and variance shifts) and simulations, disentangling the shadow costs of the policy constraints from extreme weather conditions for the lowest income stratum. I then estimate the welfare effect of zeroscaping, targeted to improve the welfare of the lowest income stratum. Finally, Section VI concludes the paper by summarizing the main findings, discussing their broader policy implications for water resource management in an era of climate change, and suggesting avenues for future research.

II. Water Utility Pricing

In this section, I introduce water utility pricing in Austin, TX, and detail the challenges water utilities generally face under volatile weather conditions. This includes the current pricing structure and the policy goals utilities set to achieve, noting in particular that the price structure is often predicated on an assumed correlation between household income and consumption levels/price elasticities. Subsequently, I present the datasets used and examine whether these assumptions could hold, and spotlight the challenges faced by the utility's pricing decision under volatile weather.

A. *Utility Pricing and Increasing Block Pricing (IBP)*

A utility company, typically a natural monopoly with large fixed costs, faces a classic challenge outlined in Ramsey's pricing problem (from Ramsey (1927)): pricing at marginal cost, though efficient, would not cover its total costs, rendering the "first-best" solution infeasible. Ramsey proposed a "second-best" solution to meet the company's financial requirements while minimizing the reduction in social welfare. This price-setting

[2] This is due to the less demanding revenue goal set by Austin Water. A different utility with a more pressing revenue-feasible requirement could also cause welfare losses for the lower-income households.

[3] An effort to remove water-intensive vegetation to save more water.

solution is known as the inverse elasticity rule, where prices are higher for consumers with lower price elasticity, as inelastic demand causes smaller reductions in consumption and thus smaller welfare distortions. In the case of a water utility, where major costs are fixed (e.g., reservoir management, purification, and distribution infrastructure) while the marginal cost per household is relatively small, Ramsey pricing helps ensure the utility's financial viability while maximizing consumer welfare.

In addition to ensuring cost recovery, the utility must also meet its policy goals. Typically, the utility must consider two major policy objectives: 1) **resource conservation** and 2) the **distributional effect** of the price. With limited natural resources, the utility is encouraged to reduce total consumption—a constraint that has become more pressing with dwindling supply caused by climate change. On the other hand, the utility must also ensure that water remains affordable for low-income consumers' basic needs. Therefore, the optimal price must strike a balance: it cannot be so high as to be unaffordable for low-income households, nor so low as to jeopardize revenue and encourage excessive consumption.

A common practical solution is to adopt a multi-part tariff with Increasing Block Pricing (IBP), which categorizes consumers into tiers based on their consumption. For all tiers, the final bill consists of a fixed charge (or access fee) to cover fixed costs and a volumetric charge (or usage fee) based on consumption to cover marginal costs. In an IBP structure, both the fixed and volumetric charges typically increase with higher tiers. For the volumetric charge, consumers pay a lower marginal price on consumption up to each kink point and a higher marginal price only on the amount exceeding it. This structure encourages high-quantity users to conserve, while lower-quantity users can benefit from a low marginal price (often below marginal cost). Predicated on the assumption that higher-quantity users typically have higher incomes and lower elasticities, IBP can be interpreted as an empirical application of Ramsey pricing that also addresses the policy goals of conservation and equity.

For this paper, I am going to specifically examine the efficacy of IBP in the context of various weather patterns. For the electricity utilities, pricing with the existing policy goals would be difficult to balance under unexpected heat waves, in addition to regular seasonal demand cycles. For the water utilities, volatile weather will cause complexities in pricing decisions as well. Urban household water usage is highly correlated to precipitation for single-family homes, where lawn watering constitutes a major portion of water consumption, far larger than essential usage like cooking, showering, and washing. Low precipitation (dry months) encourages much higher water usage, while high precipitation (rainy months) reduces usage as households may not need to use extra water for lawn watering. Therefore, for cities like Austin, where high precipitation events can occur randomly throughout the year[4], it is important to design the IBP with the policy goals of conservation and maintaining financial viability to prepare for all weather situations, and the results derived from this paper can be applied to public utilities in general, as both electricity and water utilities have similar policy goals and unpredictable demand

[4]For Austin, and many cities in Texas, high rainfall months could happen in any months between March to June, and September to November. These months could equally have low rainfall due to some years with a longer summer season.

through volatile weather.

The importance of setting the price for all weather conditions creates another price-setting challenge for utilities, as the price typically cannot change from month to month. While economists typically suggest dynamic pricing, such as that used in ride-sharing services, to combat volatile demand, in the case of water utility pricing, the price must be preset for the entire year and cannot change within a certain time period. This inability to adjust prices in response to weather stochasticity is specifically what this paper addresses.[5]

B. Water Utility Pricing in Austin, TX

For this paper, I use water utility transaction data from Austin, TX. The water utility in Austin is managed solely by Austin Water, a public entity and natural monopoly. All operations—from water supply management in Lake Travis to purification, transport, wastewater processing, and price setting—are managed by Austin Water. This means Austin Water acts as a social planner aiming to maximize welfare while pursuing policy goals, making a Ramsey pricing framework highly suitable for analysis.

Austin, and Texas in general, are known for unpredictable precipitation. A report by Nielsen-Gammon et al. (2020) noted that future precipitation trends in Texas are likely to be dominated by largely unpredictable natural variability and a projected increase in the intensity and frequency of extreme rainfall events. This makes the price-setting challenge more acute, as it is difficult to predict which months will be wet or dry, rendering preset price discrimination based on seasonality nearly impossible. In Austin, as shown in Figure A1 using NOAA data, recent precipitation has become more erratic compared to the 30-year average. This trend, combined with Austin's rapid population growth, makes residential water conservation an increasingly pressing issue.[6]

The problem faced by Austin Water is not unique. Many U.S. cities face the combined pressures of volatile weather, growing populations, and dwindling natural water supplies due to climate change.[7] This paper aims to test whether a Ramsey-style pricing solution can achieve the policy goals of conservation and distributional equity under volatile weather. I utilize a panel dataset from Austin Water for approximately $120,000$ households from May 2018 to December 2019, public data from the Travis County Appraisal District, weather data from NOAA, and high-resolution (10m $\times$ 10m) satellite imagery capturing household vegetation health. These data sources are detailed in Section III.B.

Austin Water's current pricing uses a five-tier IBP structure, detailed in Table A1. This design aims to improve distributional outcomes and encourage conservation by setting higher marginal prices and fixed payments for higher-quantity users. This is based on the assumption that higher-quantity users are correlated with higher income levels, have

[5]Some may question the reason behind this inability to implement dynamic pricing, and there are deeper reasons that need further discussion in political economy. There is simply a lack of industrial convention to deploy dynamic pricing in water utilities on a monthly basis. The unpredictability of which month will be wet or dry also makes it difficult to proactively set up a stable pricing rule (such as price discrimination by summer and winter months).

[6]According to recent census data, the Austin metro area is one of the fastest-growing regions in the U.S.

[7]The EPA confirms that hourly rainfall rates and the intensity of heavy precipitation events have increased across most of the U.S. since the 1970s.

lower price elasticities, and can afford higher prices. To better measure distributional effects, I segment households into five income strata, aligning with the five-tier price structure, to analyze how policies affect different economic groups (see Figure A.A1).

C. Descriptive and Reduced-Form Evidence

To study the impact of volatile weather on the utility's price-setting problem and the challenges its policy goals present, I will first test the assumption of a correlation between income and quantity.

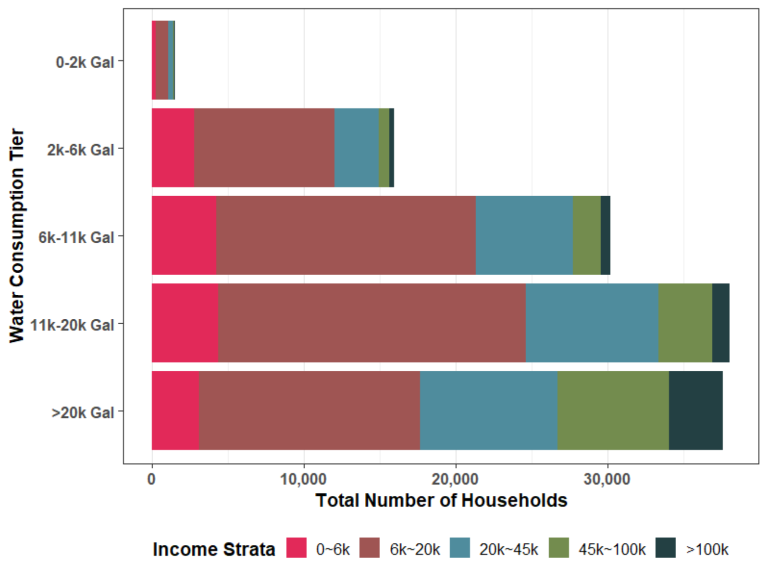

Figure 1. : Households Income Strata Composition for Each Tier

Figure 1 showcases the distributions of the average monthly quantity for each household, and the composition of the income strata for each tier. There are more households from the highest stratum that consume in the highest tier (> 20k Gal). However, more than 3000 households from the lowest strata consume on average in the highest tier, which should be mostly occupied by households from higher strata. This undermines the premise that lower marginal prices in lower tiers effectively target low-income households for distributional benefits, and this omnipresence of high-quantity consumers across all income strata poses a challenge for using IBP to achieve equity goals.

In addition, to understand the impact of weather, I examine the correlation between weather changes and quantity changes from May 2018 to December 2019. To account for seasonality, I calculated the monthly difference in precipitation (Δ Precipitation) and temperature (Δ Temperature) relative to their 30-year averages (1990-2020). Using panel data, I then calculated the percentage deviation of a household's monthly consumption from its historical average for that same month. I am particularly interested in how monthly weather deviations influence this quantity deviation and how the effect differs across income strata.

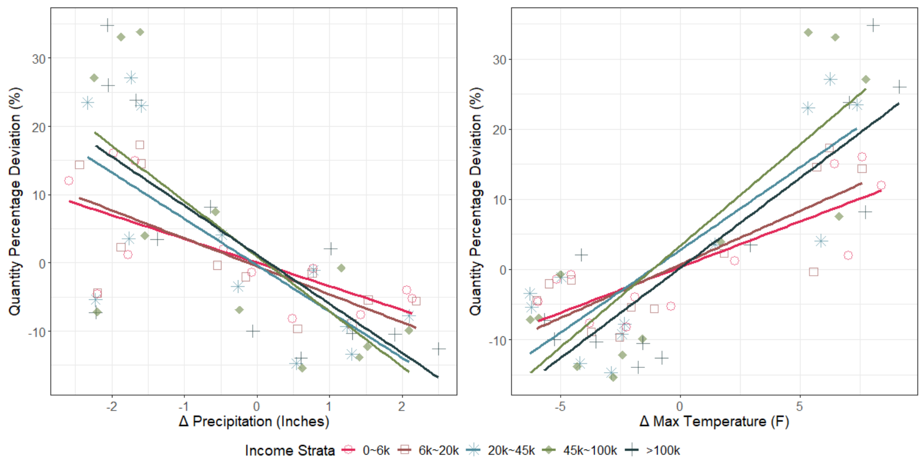

Figure 2. : Quantity Change (%) by Weather Deviation for each Income Strata

We can see that higher-than-usual precipitation typically leads to lower-than-usual consumption. However, the percentage deviation is greater for high-income households, likely reflecting higher demand for outdoor water use. A similar, though less pronounced, effect appears to exist for temperature. A reduced-form analysis of Δ Precipitation, Δ Temperature, and their interactions with income strata on quantity deviation (see Table A3) confirms these patterns.

Based on the results, higher income strata tend to have a lower baseline percentage deviation from their usual consumption (when weather differences and marginal price are zero) compared to the lowest income group. However, higher income strata (> 6k) show a stronger negative response to increases in precipitation difference. The linear response to temperature difference is less clear-cut across strata in this model; while the base effect for the reference group is positive, higher income groups show interactions that temper this positive effect, resulting in less positive linear sensitivity compared to the low-income strata. Due to this unclear effect from the reduced-form analysis, coupled with the intrinsic correlation between precipitation and single-family home water usage, I will focus on the impact of changing precipitation for the rest of the paper and observe the shifts in consumer behavior.

This heightened weather sensitivity among higher-income households is somewhat counterintuitive and challenges the assumption that they have uniformly low price elasticity. This creates potential complexities for designing an optimal pricing structure that can meet all policy goals under volatile weather. To accurately evaluate counterfactual prices and revenue risks, it is crucial to measure how behavior changes in response to price, precipitation, income, and their interactions. Therefore, a structural model is needed to incorporate the full vector of household characteristics—a necessity given the limited price variation in the dataset—followed by an empirical Ramsey model that considers weather stochasticity to estimate the optimal IBP.

III. Demand Model and Estimation

In this section, I develop the demand model for the water utility to estimate the parameters needed for optimal price calculation and counterfactual analysis. As noted earlier, residential water demand typically faces an IBP structure, meaning that the nonlinear budget constraint generated must be accounted for. Therefore, I utilize a Discrete/Continuous Choice (DCC) model estimated via Maximum Likelihood to estimate the conditional demand probabilities for each pricing tier for each household, thereby accurately estimating the demand parameters.

Each household chooses between water consumption and a numeraire good subject to its budget constraint. The pricing scheme creates a nonlinear budget constraint, and I focus exclusively on the IBP structure in this paper. I assume a log-log functional form relating demand to prices and income, which is a common assumption in the water demand estimation literature.

The first paper to introduce a demand model handling piece-wise linear budget constraints arising from price nonlinearity was Burtless and Hausman (1978) in the context of labor supply. Dubin and McFadden (1984), Hanemann (1984) and Hewitt and Hanemann (1995) laid the groundwork for applying the DCC model to residential utility demand (with Hewitt and Hanemann (1995) specifically addressing water demand). Our approach largely follows this established model structure, with minor modifications, to maintain interpretability.

A. Demand Model

The demand model assumes each household consumes both water (w) and a numeraire good (Y) (with price $= 1$), and the household's total monthly income is I. Suppose for an IBP, the marginal price for each tier k is p_k, the fixed payment is A_k, and the cutoff point between tier k and $k+1$ is q_k. Due to the nature of IBP and suppose there are a total of K tiers, $p_{k+1} \geq p_k$, $A_{k+1} \geq A_k$, $q_{k+1} \geq q_k$ $\forall k \in \{0,1,\ldots,K\}$. Together $\{p_k, q_k, A_k\}_{k=1}^{K}$ fully identify the nonlinear pricing structure. Throughout this section, the demand for every month should have a subscript of t, but for simplicity, the time-level subscript will be omitted.

Condition on the household choosing the optimal tier to be $k \leq K$ and the starting price

$p_0 = 0$, the household's budget constraint is:

(1)
$$
\begin{aligned}
I &= A_1 + p_1 w + Y \quad (\text{if } k = 1)\\
I &= A_k + p_1 q_1 + p_2(q_2 - q_1) + \cdots + p_{k-1}(q_{k-1} - q_{k-2}) + p_k(w - q_{k-1}) + Y \quad (\text{if } k > 1)\\
&= A_k + \sum_{j=1}^{k-1}(p_j - p_{j+1})q_j + p_k w + Y
\end{aligned}
$$
$$
p_k w + Y = I - A_k - \sum_{j=1}^{k-1}(p_j - p_{j+1})q_j
$$

Therefore, given the IBP pricing structure, the only way to maintain the typical utility maximization problem with a budget is to add the additional term on the income such that the Virtual Income is equal to $I + d_k$, where the correction term d_k is defined as:

$$
\begin{aligned}
d_k &= -A_1 \quad (\text{if } k = 1)\\
d_k &= -A_k - \sum_{j=1}^{k-1}(p_j - p_{j+1})q_j \quad (\text{if } k > 1)
\end{aligned} \tag{2}
$$

The concept of Virtual Income is used to correctly model consumer choice under a tiered pricing system. When a household's consumption is in a higher tier, k, the initial "blocks" of consumption (tiers 1 through $k-1$) are purchased at lower marginal prices. To simplify the analysis, we can treat the consumer's choice problem as if they face a single marginal price, p_k, for all units. For this simplified model to yield the same utility-maximizing choice, the consumer's income must be adjusted upward to account for the savings on the initial, lower-priced blocks.[8]To visually illustrate, consider a simpler model with 2 tiers without loss of generality:

Since consuming at a higher tier will grant paying the lower quantity portion at a lower marginal price, the correction term d_k (which will be > 0 in the higher tier given that $p_k \leq p_{k+1}\ \forall k$ and A_k is not too large) has an income effect on the overall budget set. This correction term will allow us to solve for the utility maximization problem using the marginal price for each tier without causing any trouble.

Define the utility function for the household between water and numeraire good to be $U(w, I)$. Condition on tier k being the optimal choice for the household, define the conditional water demand function to be $g(p_k, I)$, where the functional form will be introduced later. Then the conditional indirect utility function for tier k is:

$$
\begin{aligned}
V(p_k, I) &= \max_w U(w, I - p_k w)\\
&= U(g(p_k, I), I - p_k g(p_k, I))
\end{aligned} \tag{3}
$$

[8]This additional income, often denoted as d_k, can be considered an income effect. It allows the household to be modeled as if it faces the marginal price p_k across all quantities, even though the initial blocks were purchased more cheaply.

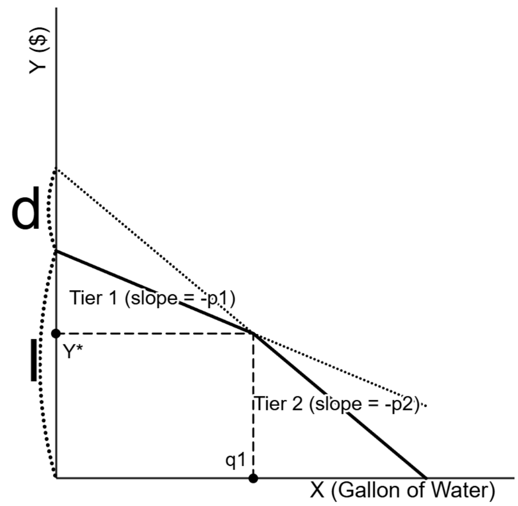

Figure 3. : Virtual Income $I+d_2$ when $K=2$

For general U and budget sets, there will be a case where the household will have multiple optimal tiers such that the conditional indirect utility function is not unique. However, fortunately, given the case of IBP, Hausman (1979)'s theorem showed that if the budget set is convex, the optimal tier will be unique. In addition, the concavity of the utility function U will make sure the optimal consumption point $g(p_k,I)$ resides in the specific quantity boundaries of each tier. Therefore, without loss of generality, the unconditional indirect utility function for the total tier $K=2$ is defined as:

$$V(p,I) = \begin{cases} V(p_1,I) & \text{if } q_1 \geq g(p_1,I) \\ V(p_2,I+d_2) & \text{if } q_1 < g(p_2,I+d_2) \\ U(q_1,I-A_2-p_1q_1) & \text{if } g(p_2,I+d_2) \leq q_1 < g(p_1,I) \end{cases} \tag{4}$$

The form of the unconditional indirect utility function can be expanded towards a more general case. I would like to highlight the case of $g(p_2,I+d_2) \leq q_1 < g(p_1,I)$ to make sure the Incentive Compatibility constraints are satisfied throughout the entire quantity line. The reason for the inclusion of this case is explained in detail in Section A.A2.

So far, I have covered the decision process of the household's water demand, in which they solve for a utility maximization problem and solve for $g(p_k,I)$ for all k, and choose the k that maximizes their utilities only with the assumption that the demand function is concave without imposing any functional form on the demand function. A typical functional form used by recent literature of water demand is a log-log demand function[9],

[9] A couple of relatively recent examples of using this functional form for water demand are Hewitt and Hanemann (1995), Olmstead, Hanemann and Stavins (2007), and Wolak (2016). Others, like Szabo (2015) used a linear demand function.

I will use the same approach as the data of water quantity w_i is very right-skewed with only positive values. Therefore, a better parametric assumption for the distribution of w_i should be log normal, which will induce a log-log demand function. With vectors of household characteristics X and weather Z, the conditional indirect utility function for tier k has the functional form of:

$$V(p_k, I+d_k, \beta) = -\exp(\beta_1' X + \beta_2' Z + c + \varepsilon)\frac{p_k^{1-\alpha}}{1-\alpha} + \frac{(I+d_k)^{1-\rho}}{1-\rho} \tag{5}$$

This functional form, combined with Roy's Identity, will give us the log-log demand form for tier k:

$$\log(g(p_k, I+d_k)) = \log(w_k) = \beta_1' X + \beta_2' Z - \alpha \log p_k + \rho \log(I+d_k) + c \tag{6}$$

where d_k is defined in Equation 2, α is the log price effect, ρ is the log virtual income effect, and c is the constant. The observed log consumption includes unobservables ε: $\log(w) = \log(w_k) + \varepsilon$. Following the same parametric assumption of Olmstead, Hanemann and Stavins (2007), the unobservables have two terms $\varepsilon = \eta + \nu$. $\eta \sim N(0, \sigma_\eta^2)$ represents the household-level heterogeneity observed by households (e.g., preferences), while $\nu \sim N(0, \sigma_\nu^2)$ represents the household-level perception/optimization error that is unobserved by households even ex-post. The idea is that it is nearly impossible for a household to precisely consume the water quantity (w_k) they would ideally demand given their preferences. For instance, there will always be extra cold water down the drain when she demands warm and hot water, for example. For all households, they are risk-neutral with $E[\nu] = 0 \ \forall k$. Neither the error terms are observed by the econometrician, hence the parametric assumption. In addition, this parametric assumption can allow us to derive a closed-form likelihood function for MLE. Given the structure of the error term, the unconditional ex post water demand after the realization of the error term is:

$$\log(w) = \begin{cases} \log(w_1) + \eta + \nu & \text{if } \eta \leq \log(q_1) - \log(w_1) \\ \log(q_1) + \nu & \text{if } \log(q_1) - \log(w_1) < \eta \leq \log(q_1) - \log(w_2) \\ \log(w_2) + \eta + \nu & \text{if } \log(q_1) - \log(w_2) < \eta \leq \log(q_2) - \log(w_2) \\ \log(q_2) + \nu & \text{if } \log(q_2) - \log(w_2) < \eta \leq \log(q_2) - \log(w_3) \\ \dots & \\ \log(q_{K-1}) + \nu & \text{if } \log(q_{K-1}) - \log(w_{K-1}) < \eta \leq \log(q_{K-1}) - \log(w_K) \\ \log(w_K) + \eta + \nu & \text{if } \log(q_{K-1}) - \log(w_K) < \eta \end{cases} \tag{7}$$

and the likelihood function for the observed water demand w_i for household i (omitting i

subscript below for brevity) given the parametric assumption of η, ν is:

(8)

$$f(w_i|X,Z) = \sum_{k=1}^{K}\left[\frac{1}{\sqrt{\sigma_\eta^2+\sigma_\nu^2}}\phi(s_k)(\Phi(r_k)-\Phi(n_k)) + \frac{1}{\sigma_\nu}\phi(u_k)(\Phi(m_k)-\Phi(t_k))\right]$$

where

$$t_k = (\log(q_k) - \log(w_k))/\sigma_\eta$$

$$r_k = (t_k\sigma_\eta - \rho_s s_k\sqrt{\sigma_\eta^2+\sigma_\nu^2})/(\sigma_\eta\sqrt{1-\rho_s^2})$$

$$\rho_s = \sigma_\eta/\sqrt{\sigma_\eta^2+\sigma_\nu^2}$$

$$s_k = (\log(w_i) - \log(w_k))/\sqrt{\sigma_\eta^2+\sigma_\nu^2}$$

$$n_k = (m_{k-1}\sigma_\eta - \rho_s s_k\sqrt{\sigma_\eta^2+\sigma_\nu^2})/(\sigma_\eta\sqrt{1-\rho_s^2})$$

$$m_k = (\log(q_k) - \log(w_{k+1}))/\sigma_\eta$$

$$u_k = (\log(w_i) - \log(q_k))/\sigma_\nu$$

where Φ is standard normal cdf and ϕ is standard normal pdf.

and $q_0 = 0, q_K = \infty, w_k = g(p_k, I + d_k)$ is defined implicitly via Equation 6

Empirically, this likelihood function is used to calculate the likelihood of the observed water consumption w_i given the model parameters $(\alpha, \rho, \beta_1', \beta_2', c, \sigma_\eta, \sigma_\nu)$ and data $(X_i, Z_i, p_k, q_k, A_k, I_i)$.[10] The derivation of the likelihood function is explained in a 2-tier case in Section A.A2 without loss of generality.

B. Empirical Model and Data

In this subsection, I outline the empirical model, based on the demand function (Equation 6) and the likelihood function (Equation 8), and describe the data used for maximum likelihood estimation.

The core dataset consists of panel data detailing monthly water transactions for single-family households in Austin, TX, from 2018-2019. These data include payments, total water usage, the billing date for each month, and the household's longitude and latitude. The data were provided by Austin Water, the public utility monopoly responsible for water services in Austin.

To operationalize the demand function, I incorporate supplemental datasets. These include two variables in the household characteristics matrix (X): the number of bathrooms per household and the time-variant household Normalized Difference Vegetation Index (NDVI). The number of bathrooms serves as a proxy for indoor water usage, while NDVI

[10] In the actual empirical model, I increase the flexibility by adding interaction term of price effect and income effect, therefore α is a function depends on parameter β_3', β_4', c_α and data X_α, Z_α, and ρ is a function depends on parameter β_5', c_ρ and data X_ρ, Z_ρ. The details will be discussed in Section III.B.

accounts for outdoor water usage, which is typically larger in volume and more sensitive to weather variations. The application of NDVI to water demand analysis was introduced by Wolak (2016). It aims to approximate lawn watering habits, as higher NDVI values indicate healthier vegetation, suggesting greater watering efforts by the household. Details on constructing and interpreting NDVI can be found in Section A.A2. For the time-variant weather matrix (Z), I include the monthly average maximum daily temperature, the monthly interquartile range (IQR) of maximum daily temperature, total monthly precipitation, and the monthly IQR of total precipitation.[11]

To enhance model flexibility, I allow the log price effect (α) and the log virtual income effect (ρ) to depend on household characteristics and weather conditions, thereby generating interaction terms. Specifically, α depends on the number of bedrooms, household NDVI, monthly average maximum daily temperature, and total monthly precipitation. ρ depends on the number of bedrooms, NDVI, and the number of heavy water-use appliances (including pools, hot tubs, sprinkler systems, fountains, and car washes).[12]

All other household characteristics, except for NDVI, were collected from public data provided by the Travis County Appraisal District (TCAD). I use data from 2018 to match the time frame of the transaction data. The TCAD data also include addresses, which were matched via geo-spatial analysis with the longitude and latitude data from Austin Water. I also collected lot size data from TCAD. Weather data were collected from the National Oceanic and Atmospheric Administration (NOAA), using data from approximately 120 weather stations in Austin. Since NDVI and weather data are collected on a calendar month basis, while household water billing cycles do not typically align with calendar months, I prorate both the weather and NDVI data to match each household's billing cycle.[13] Household income data were estimated using zipcode-level average homeowner income data from the IRS combined with normalized household value data from TCAD, under the assumption that the normalized variance of household values within each zipcode mirrors the variance of household income.

After matching all datasets and eliminating outliers, the final sample consists of $127,323$ households with $2,345,742$ transaction records. Summary statistics can be found in Section A.A2.

C. Estimation Result and Identification

After conducting the maximum likelihood estimation using Equation 8 with $k = 5$ and the data discussed in Section III.B, the final estimation results can be found in Table A5.

With the panel data, both time-variant and time-invariant variables provide significant heterogeneity among households and across different months. This helps identify the parameters associated with household characteristics, income, and their interaction terms. Moreover, for weather variables, instead of using weather data for the entire city, I found

[11]Note that both IQRs account for the spread of the data within each month, which differs slightly from what I mean by growing weather variance, namely the increasing variance of weather patterns within a year.

[12]In practice, I used $\alpha = \exp(\beta_3' X_\alpha + \beta_4' Z_\alpha + c_\alpha)$, and $\rho = \beta_5' X_\rho + c_\rho$ to maximize the bias-variance trade-off of the estimation results. Regardless of the functional form, the expansion of both α and ρ is to provide flexibility in identifying heterogeneity of price effect and virtual income effect of the households, through the interaction terms.

[13]This practice was first introduced by Train et al. (1984).

the closest weather station for each household to introduce small variations, providing heterogeneity to help identify the weather parameters. Nevertheless, identifying the price effect and its interaction terms presents a challenge, as there is little or no price variation (neither through time nor geography) in the data. Another concern highlighted in past literature (e.g., Borenstein (2009); Ito (2014)) is that households may not react to the observed marginal price, as prices and the total payment amount are typically observed only after the billing cycle ends; instead, they might respond to the average price they face.

Fortunately, both of these identification concerns regarding price parameters in the DCC model have been addressed by Olmstead (2009). She demonstrated that each household, for each month, is optimizing over the entire price schedule. Consequently, the econometrician can recover the parameter estimates, along with the probabilities that households consume on each of their budget segments and at each kink point, directly from the DCC model results. In addition, the DCC model is estimated without ever determining the 'observed' marginal price of consumption – all the prices and the kink points enter the likelihood function, regardless of where consumption is actually observed. Olmstead (2009) also pointed out that price elasticity estimation has smaller bias when using the DCC model if variation in demand is driven primarily by household preferences (η) instead of the perception error (ν), and our estimation results align with this requirement, showing that σ_η is much larger than σ_ν. Furthermore, the model generates less bias if the existing price jumps between different tiers are sufficiently salient. Although there is no definite conclusion on how salient the price jump is needed to minimize the estimation bias, the marginal price differences between tiers in Austin are among the largest in the US[14].

To further address the price elasticity identification challenge, I also utilize data from a small category of consumers enrolled in the Consumer Assistance Program (CAP). Based on certain income requirements, these consumers enjoy lower marginal prices (see Section A1). Although the variation is small, these consumers provide a small degree of price variation to aid in the identification of price elasticity.

Another potential endogeneity issue involves the use of NDVI, as the NDVI of a given calendar month could be the result of water usage in that same period. Therefore, I used a lagged term for NDVI, meaning that water demand this month is dependent on the previous month's NDVI. The idea is that households decide how much water to consume this month based on the health of their lawn vegetation last month. A higher previous month's NDVI may indicate the household takes more care of its lawn, helping to account for variations in outdoor water usage.

Based on the identification strategy discussed above, I now discuss the price and income elasticity derived from the DCC model. Due to the nonlinear pricing structure, I cannot directly use α and ρ to calculate price and income elasticities. Instead, by following Olmstead, Hanemann and Stavins (2007), I use a simulation-based approach. For a small number ξ, I calculate how expected demand changes in response to a small change

[14]See `https://bseacd.org/conservation-based-rate-structures/` for "Conservation-Based Rate Structures" - Barton Springs/Edwards Aquifer Conservation District

in the status quo marginal price (p_0), holding income (I) and other factors constant.[15] The formula for price elasticity is as follows, where the same structure applies to income elasticity:

$$\text{(9)} \qquad \text{Price Elasticity: } (E[g((1+\xi)p_0, I)] - E[g(p_0, I)])/(\xi E[g(p_0, I)])$$

The median price elasticity is estimated to be -0.395. Compared to previous literature that uses the structural DCC model for Increasing Block Pricing (IBP) in developed countries, our estimate is within the range of Pint (1999)'s estimations ($-1.24 \sim -0.04$) and more strictly, Olmstead, Hanemann and Stavins (2007)'s estimations ($-0.59 \sim -0.33$). Our estimate is more inelastic compared to Hewitt and Hanemann (1995)'s ($-1.63 \sim -1.57$). Our estimate is also close to that of Olmstead (2009) (reporting -0.609 for a specific DCC model specification, using data from 11 cities with diverse weather patterns) and Strand and Walker (2005) ($-0.3 \sim -0.1$, using data from 17 cities in Central America, where water demand was primarily for indoor usage). The median income elasticity is estimated to be 0.112, which is close to estimations from previous literature for developed countries, such as Olmstead, Hanemann and Stavins (2007) ($0.1786 \sim 0.1865$) and Olmstead (2009) (0.1865).

To further explore the heterogeneities of price elasticities across income strata, I plot the empirical density of the price elasticities for all 5 income strata.

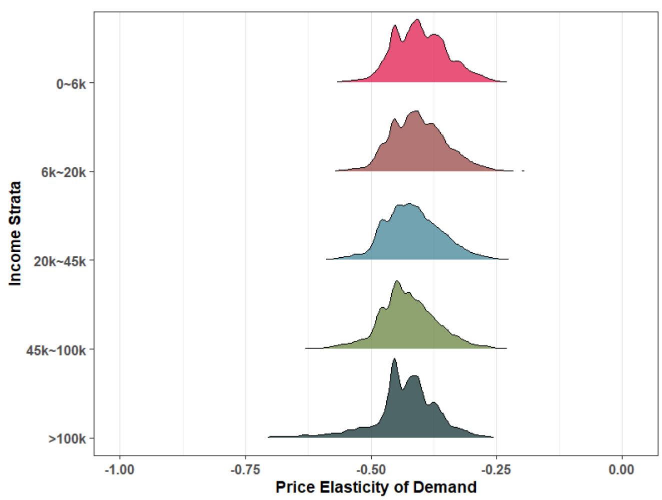

Figure 4. : Price Elasticity Density by Income Strata

There are very few heterogeneities of price elasticity across income strata, showcasing

[15]Additionally, due to the flexibility of the price elasticity function, I limit the sample for empirical analysis to the year 2019, as opposed to from demand estimation, I used data from Jun. 2018 to Dec. 2019. This approach helps avoid overestimating the more rainy summer months of 2018.

that the assumption that higher income households with higher consumption will have low elasticities is not correct, breaking the premise of using IBP to achieve the distributional goal. On the contrary, the highest income stratum even has a larger left tail, meaning the lower end of the highest income stratum is even more elastic compared to other strata. To further explain the reduced-form evidence from Figure 2, I separate the data into wet and dry months based on the median precipitation value.

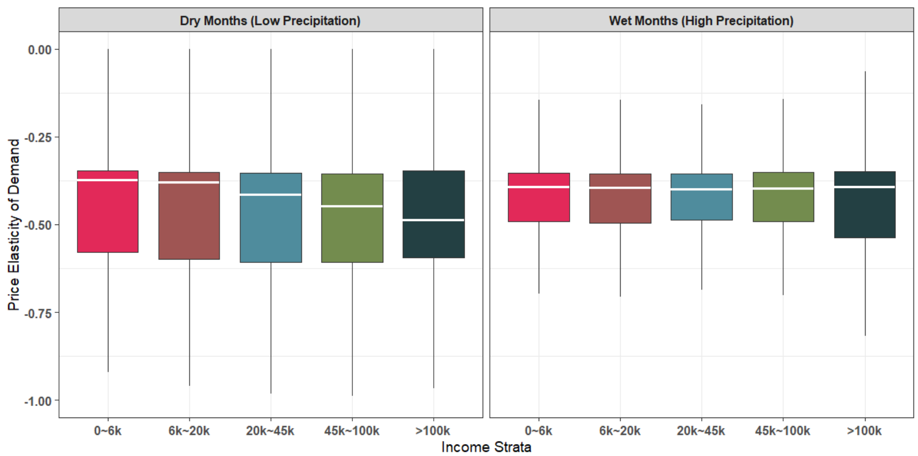

Figure 5. : Price Elasticity Boxplots by Income Strata - Wet vs. Dry Months

There is a clear trend in the dry months where the median price elasticity is lower for higher income strata (larger in magnitude)[16], meaning households with higher income are more responsive to price changes under low precipitation, while lower income strata do not share this difference. This counterintuitive result is likely due to the interaction between price effect and precipitation: when it rains a lot, few households across income levels have excess outdoor water consumption that responds to price, but when it is dry, households with the preferences of greener lawns and larger pool usage are more responsive to price change, and these households tend to have higher income. In addition, higher-income households are more likely to be equipped with better technology to precisely control their water usage [17], which will also contribute to higher price elasticities. From Figure A5, we can see that 14.4% of the highest income stratum households are more elastic (lower than the overall median), and consume on average more than 20k gallons.

This observation of high elasticities for higher-income households under dry weather, combined with the omnipresence of high quantity consumers across all income strata (see Figure 1), adds extra challenges to the optimal price design, especially in dry months. Under the inverse elasticity rule, the price should decrease for these households. How-

[16]This fits the reduced-form evidence from Figure 2.
[17]Like smart meter, smart sprinkler, etc.

ever, a price decrease would result in an overblown increase in the consumption of the high-income, more elastic, and high quantity consumers, which poses a serious issue in reaching the conservation goal of the utility. In addition, a price decrease for high-income households will certainly create complexities in reaching the distributional goal of the optimal price. More discussions regarding the price elasticity heterogeneities among income strata can be found in Section A.A2.

D. How weather shifts the Demand Curve?

From the structural demand model, I have estimated the heterogeneity of price elasticities across income strata and weather conditions, and observed that under dry weather, the higher income households would become more elastic, which creates complexities under Ramsey's inverse elasticity rule. This also showcases the importance of precipitation in the discussion of water utility pricing. In this subsection, I will introduce four alternative weather conditions and observe the changes in quantity and payments under the status quo pricing scheme to understand how weather shifts the demand curve. The four weather conditions are: 1) dry weather, where the precipitation decreases by 0.25 inches for every month compared to the status quo weather, 2) rainy weather, where the precipitation increases by 0.25 inches for every month compared to the status quo weather, 3) low variance, where the new precipitation standard deviation is $0.75\times$ the status quo standard deviation, and 4) high variance, where the new precipitation standard deviation is $1.25\times$ the status quo standard deviation. I only changed precipitation and kept all other weather variables unchanged. In addition, all other house characteristics remain the same as their status quo counterparts. I will then compare the new quantities and payments through the alternative weather and compare them to the status quo. I will present the overall distribution of the data through empirical cdf and then focus the data to the lowest income stratum ($0 \sim 6$k) and the highest income stratum (> 100 k).

When the weather becomes drier, less precipitation will shift the demand curve rightward, causing quantity and payments to increase. Conversely, when the weather becomes more rainy, more precipitation will shift the demand curve leftward, causing quantity and payments to decrease.[18]This means that in order to achieve a better conservation goal for the utilities, the status quo price is too low, at least in the higher tier, to better curb the excessive demand under dry weather conditions, while the financial feasibility concern is minimized. On the other hand, under rainy weather, the conservation goal is easy to achieve, while financial feasibility concerns are more pressing.

When the weather decreases in variance, the demand distribution will have smaller spread as well, meaning consumptions left of the median will increase, and right of the median will decrease. However, since both the precipitation, and quantity are very right-skewed, causing the decrease of the consumptions larger than mean more dominant. Therefore, the empirical cdf largely resembles the pattern of more rainy weather, where the financial feasibility concerns are more pressing. Conversely, when the weather increases in variance, the demand distribution will have larger spread, causing the con-

[18]The "bumps" in either empirical cdfs, are due to the consumer bunching around the kink points.

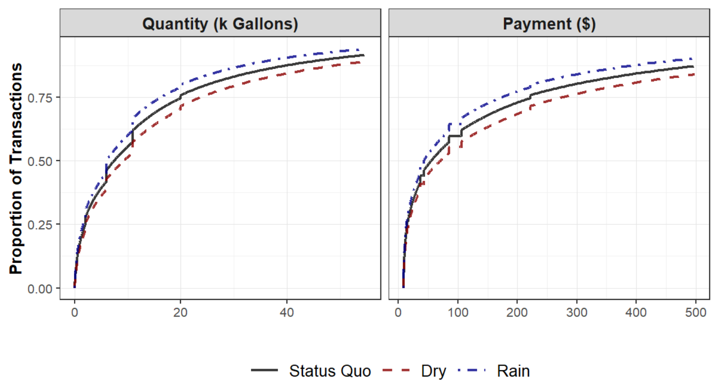

Figure 6. : Changes in Quantity and Payments Distributions between Weather Conditions - Dry and Rainy

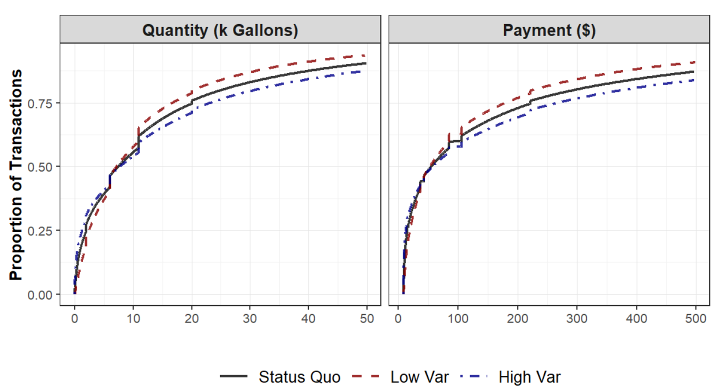

Figure 7. : Changes in Quantity and Payments Distributions between Weather Conditions - Low and High Variance

sumption right of the median to increase. These shifts on the demand curve are more dominant due to the same reason, making the empirical cdf largely resembles the pattern of drier weather, where the conservation goal is harder to meet.

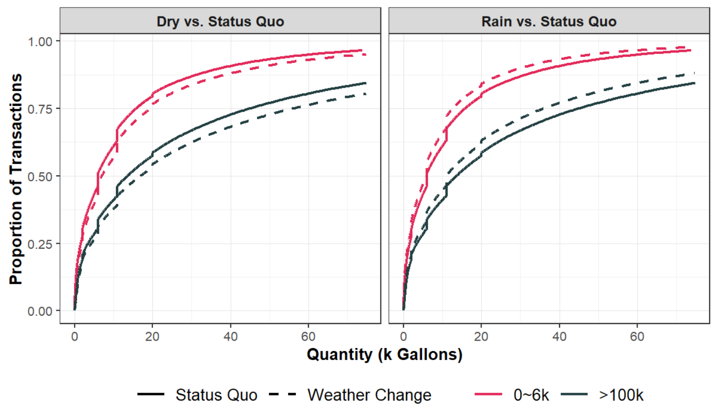

Figure 8. : Changes in Quantity Distributions between Lowest and Highest Income Strata - Dry and Rainy

When specifically comparing the demand distribution between income strata, I observe the difference between the empirical cdf from new weather pattern and the empirical cdf of the status quo weather, and the highest income stratum has larger differences, for both dry and rain weather conditions. This showcases the "weather elasticity" is larger in magnitude for the higher income households, which aligns with the evidence I saw from Figure 2. Low and high variance weather conditions show similar patterns. The low elasticities of the low income households will be quite concerning regarding the distributional effect, especially when the demand curve shift rightwards. When the demand curve shifts rightwards, the conservation goal will require the price to increase at least at higher tiers. Balancing both goals with a single price will be challenging for the utilities and will certainly create a shadow cost, which will inevitably decrease welfare. Based on the counter-intuitive price elasticity results in Figure 4 and the existence of high quantity households in lower income strata in Figure A5, it is reasonable to question if the price itself can achieve both policy goals, as well as improve the distributional effect. In the next section, I will formally define the empirical optimal price model and further implications and caveats of the challenge created by price elasticities interacted with weather across different income strata can be found in Section A.A2.

IV. Ramsey Pricing Model

In this section, I establish the theoretical basis for the Ramsey pricing problem (from Ramsey (1927)) faced by a natural monopoly water utility company with the policy constraints it is facing. Finally, I introduce the empirical model used to estimate the optimal

Ramsey pricing under different extremes of weather conditions.

As previously established, Ramsey proposed the "second-best" solution to address a company's financial requirements while minimizing the reduction in overall social welfare, employing the inverse elasticity rule. Water utility companies, particularly in urban metropolitan areas with high fixed costs, often adopt the Ramsey format of IBP within a multi-part tariff structure to meet financial requirements and policy goals of conservation and equity. For simplicity, when discussing the theoretical basis for the Ramsey problem, I treat the entire pricing structure as a single price p, as opposed to the pricing vector of IBP in a multi-part tariff. In the empirical model, I reintroduce the multi-part IBP to derive more realistic solutions.

The theoretical Ramsey model, introduced by Ramsey (1927), is extended to include additional policy constraints in the generalized Ramsey model, as discussed in literature such as Coady and Drèze (2002). For the empirical model, I largely follow the setup of the conservation constraint for the Ramsey problem as presented in Wolak (2016). To empirically measure consumer welfare, I adopt the framework and definition of equivalence variation for nonlinear budget constraints from Hausman (1981), Reiss and White (2005), and Ruijs (2009).

A. Ramsey Pricing Model with Conservation Constraint

In this subsection, I briefly showcase the changes from the classic Ramsey model by using a simplified pricing model with the price p, chose by the water utulity to maximize consumer welfare, subject to two constraints: 1) annual revenue ($R(p)$) is greater than or equal to an exogenous total annual cost (C) that is independent of demand (simplifying for large fixed costs),[19] and 2) total annual quantity ($q(p)$) is less than or equal to an exogenous annual quantity upper bound ($\bar{Q}$).

$$\begin{aligned} \max_{p} \quad & CS(p) \\ \text{s.t. } & R(p) - C \geq 0 \\ \text{and } & q(p) \leq \bar{Q} \end{aligned} \tag{10}$$

Solving this Lagrangian yields the following result:

$$\begin{aligned} \underbrace{\frac{p-C}{p}}_{\text{Markup}} &= \underbrace{\frac{\lambda-1}{\lambda}\frac{1}{\varepsilon}}_{\text{Ramsey Rule}} + \underbrace{\frac{\mu}{\lambda p}}_{\text{Conservation Penalty}} \\ p^* &= \frac{C+\mu/\lambda}{1-\frac{\lambda-1}{\lambda}\frac{1}{\varepsilon}} \end{aligned} \tag{11}$$

The classic Ramsey rule states that the markup is proportional to the inverse of price elasticity (ε), but the optimal markup here includes an additional term that serves as a

[19]In reality, the total cost of maintaining urban water is not related to demand fluctuation; regardless of the amount of water consumed, maintenance and transportation fees do not change, making this a reasonable assumption.

penalty for the conservation constraint. The optimal price depends on λ (the Lagrangian multiplier for the revenue constraint), μ (the Lagrangian multiplier for the conservation constraint), the exogenous cost C, and the price elasticity ε. If μ increases, the shadow price of the conservation constraint increases, and the utility has a greater incentive to raise the markup. If $\mu = 0$, the conservation constraint is not binding, and the standard Ramsey rule is recovered. If C increases, the price will need to increase. If ε increases, demand becomes more elastic, and the price will need to decrease, which is consistent with the standard Ramsey rule. If I denote $k = \frac{\lambda-1}{\lambda}$, the price can be expressed as $p = \frac{C+\mu(1-k)}{1-\frac{k}{\varepsilon}} = \frac{C+\mu}{1-\frac{k}{\varepsilon}} - \frac{1}{k-\frac{1}{\varepsilon}}$. If λ increases, k (which is less than 1) will also increase. The first term increases. The second term, being minus a negative number since $k - \frac{1}{\varepsilon} < 0$, also increases. This means that when λ increases, the price will increase.

B. *Empirical Model*

In this subsection, I develop the empirical model used to estimate the optimal price derived from Equation 10. The empirical model's goal is to extend the Ramsey model into a more realistic setting. [20] To keep consistency with the demand model, the optimal price ($\mathbf{p}$) consists of three vectors: marginal prices ($\{p_k\}_{k=1}^5$), kink points ($\{q_k\}_{k=1}^4$), and fixed payments ($\{A_k\}_{k=1}^5$). To maintain Austin Water's current pricing structure of 5 tiers with 4 kink points, $\mathbf{p}$ has a total of 14 parameters, and I assume the pricing structure does not change[21]. The optimal price, which depends on weather Z, maximizes the annual total equivalence variation (EV) weighted by household income I to boost the distributional effect.[22] To account for weather-level stochasticity, ideally, I would need to utilize the real distributions of weather to derive E_Z. However, it is challenging to estimate this real distribution of weather consistently by using either data from very recent years (which lack sufficient data points for empirical distribution derivation) or long-term data (which are affected by climate change trends). Instead, I assume the utility faces an expectation based on its prior[23], and then calculates the optimal price based on its weather expectation. Later, I will perform a robustness check by adding a Monte-Carlo disturbance to simulate weather prediction error and generate a distribution of Z. Further details on simulating the distribution of Z can be found in Section A.A3. Aside from stochasticity from Z, the demand estimation using MLE (Equation 8) also introduces household-month-level demand stochasticity from ε_h, as the econometrician observes neither η_h nor ν_h. I use Monte-Carlo simulations for both to generate their

[20]The model is partially based on Wolak (2016).

[21]This includes the assumption that each marginal price and fixed payment will be non-decreasing in tiers.

[22]As pointed out by Feldstein (1972), for public pricing to account for distributional equity, the price setting should set the welfare weight according to the social marginal utility of income. In addition, a logarithmic social utility of inequality and $\varepsilon = 1$ for the Atkinson Index (Atkinson (1970)) result in a weighting proportional to 1/I.

[23]In reality, utilities would form a general trend of next year's weather based on phases of the El Niño-Southern Oscillation (ENSO). For example, next year's weather will be drier in general, causing precipitation to be lower in the mean.

distributions.[24]

$$\begin{aligned} \mathbf{p}^*(Z) = \underset{(\mathbf{p},q,A)}{\arg\max} \sum_h \left[\Big[w_h \cdot EV_h(Z;\mathbf{p},\mathbf{p}_0,I) \Big] - \lambda \cdot \max(0, C - R_h(Z;\mathbf{p})) \right] \\ \text{s.t.} \quad P\left(\sum_h q_h(Z_h;\mathbf{p}) \leq \bar{Q}) \right) \geq 0.95 \end{aligned} \tag{12}$$

Note that for the two exogenous thresholds: $C = \sum R(Z_0;\mathbf{p}_0)$ and $\bar{Q} = \sum q(Z_0;\mathbf{p}_0)$. Both Z_0 and $\mathbf{p}_0$ represent the status quo weather and price. This means I compare the counterfactual revenue and quantity to their status quo counterparts. In particular, due to the lack of detailed information on costs, I use the status quo revenue as the benchmark to evaluate financial viability. Utilities usually ensure their revenue is just enough to cover the annual cost to avoid excessive welfare distortion, which makes this a valid assumption. I further discuss the validity of this assumption in this case in Section A.A3. For the conservation constraint, I introduce chance-constrained programming to provide more flexibility in price setting. It essentially requires that the probability of the counterfactual annual total quantity for all households being smaller than the status quo annual total quantity is greater than 0.95. The structures of both constraints are adopted from Wolak (2016). Unlike the typical Ramsey pricing setting, I set up the revenue constraint as a revenue loss term $-\lambda \cdot \max(0, \text{loss})$, such that when the revenue requirement is not met, it incurs a cost to the entire economy, but does not necessarily award any positive value when the requirement is met. The goal is to capture the utilities' surcharge mechanisms or debt-service coverage rules that make large revenue losses more costly. The parameter λ controls the weight between the monetary value of welfare, measured by the weighted equivalent variation, and the loss of revenue for the utilities. Mathematically, it is similar to the Ramsey model, but empirically, this provides more flexibility as, instead of a hard constraint, the utilities are allowed to have some annual losses.

The weight in front of the welfare, $w_h = I_{\text{median}}/I_h$ is to boost the distributional effect from the pricing optimization procedure. I took the inverse of the household income and centered the median to be 1, such that the welfare part and the revenue part are comparable. The weight is an assumption on the policy of the utility, as they will actively try to boost the distributional effect from the pricing. By making it such that higher than median income will receive smaller weights ($w_h < 1$) and lower than median income will receive larger weights ($w_h > 1$), the aggregated welfare will, through the optimization process, actively look for a price that can boost distributional effects.[25]

The weight of the revenue loss λ is calculated by selecting a grid of $\lambda = [0.25, 0.5, 0.75, 1, 1.5, 2, 5, 10]$, and solving for the optimal price under status quo weather. I find that when $\lambda = 0.5$, it generates the result closest to the status quo prices. No further fine-

[24] I will generate η_h on the household level to avoid overfitting.

[25] The idea of using income inverse as the base for social welfare calculation was first developed by Atkinson (1970). When $\varepsilon = 1$, the Atkinson index is $1/I$. The normalization of the weights such that a household with the median income receives a weight of 1 is a standard approach in applied cost-benefit analysis (e.g., Office of Management and Budget, 2023).

grained grid search has been implemented, as I largely need to know the policy preference of the utility. $\lambda = 0.5$ means they weight the consumer welfare twice as important compared to revenue losses, which is reasonable for a social planner.

The welfare itself is defined as the equivalence variation ($EV_h(Z;\mathbf{p},\mathbf{p}_0,I)$). It is calculated by obtaining the expenditure function from the indirect utility function and then calculating the welfare effects from the price change. The framework is developed by Hausman (1981) and Reiss and White (2005), and the formal definition is from Ruijs (2009).

Based on the indirect utility function $V(\mathbf{p}, I + d_k)$ from Equation 3, the expenditure function $e(\mathbf{p}, u)$ for a household choosing tier k is:

$$e(p_k, u) = \left[(1-\rho)\left(u + \exp(\beta_1' X + \beta_2' Z + c + \varepsilon)\frac{p_k^{1-\alpha}}{1-\alpha}\right)\right]^{\frac{1}{1-\rho}} \tag{13}$$

If a household's counterfactual demand falls within tier k, rather than at a kink point q_k, its equivalent variation is:

$$EV(p_0, p, I) = e(p, V'(p)) - d_k^0 - I \tag{14}$$

where $V'(p)$ is the new counterfactual utility generated by the new price p, and d_k^0 is the virtual income correction term from the original price scheme, i.e., $d_k = -A_k - \sum_{j=1}^{k-1}(p_j - p_{j+1})q_j$. The equivalent variations are adjusted for the amount of subsidies received due to the nonlinearities in the budget set. Note that the expenditure function ($e(p_k, u)$) is used to solve for the virtual income from the indirect utility function, instead of I, hence there is no correction term d_k here. If the price does not change, $EV(p_0, p_0, I) = 0$ and $e(V_0, p_0) = I + d_k^0$.

On the other hand, when the predicted consumption is at a kink point q_k, the above definition does not apply as these consumers are not technically facing the new marginal price from their chosen tier.[26] The equivalent variation is generated from $\bar{p}$, where $\bar{p}$ is the price at which the demand function (from Equation 6) generates the result q_k. If I denote $\mathcal{A} = \beta_1 X + \beta_2 Z + c + \varepsilon$, and $\bar{V}I$ denotes the corresponding virtual income, the idea is to solve for $(\bar{p}, \bar{V}I)$ as a substitution effect where $(\bar{p}, \bar{V}I)$ generates the same utility. This means: $\begin{cases} V(p_k, I + d_k^0) & = V(\bar{p}, \bar{V}I) \\ \log(q_k) & = \mathcal{A} - \alpha\log(\bar{p}) + \rho\log(\bar{V}I) \end{cases}$. I can solve for $\bar{p}$:

$$\log(q_k) = \mathcal{A} - \alpha\log(\bar{p}) + \frac{\rho}{1-\rho}\log[\exp(\mathcal{A})\frac{1-\rho}{1-\alpha}(\bar{p}^{1-\alpha} - p_k^{1-\alpha}) + (I + d_k^0)^{1-\rho}]$$

For this demand function and indirect utility function, I need to solve this nonlinear function numerically to obtain $\bar{p}$. Then, the equivalent variation for predicted demand at

[26]To see more details of the additional case, see Section A.A2.

the kink point q_k is:

$$EV(p_0, p, I) = e(\bar{p}, V'(p)) - (\bar{p} - p_0)q_k^0 - d_k^0 - I \tag{15}$$

V. Counterfactual Analysis

In this section, I analyze the welfare effects of counterfactual optimal prices under various precipitation patterns, focusing on the differences in these effects across income strata. I quantify the associated welfare by comparing the new and status quo values for revenue ($R\%$), quantity ($Q\%$), and the equivalent variation to income ratio (EV/I). These measures capture the changes from both the price structure (including its policy constraints) and the weather itself. The results indicate that the lowest-income stratum is the most susceptible to welfare decreases from both exogenous weather shocks and the policy constraints within the price optimization process. I then disentangle the welfare change from these two sources to calculate the shadow cost of the policy constraints on the lowest-income stratum. Lastly, I propose zeroscaping/xeriscaping as a policy solution to specifically improve the distributional outcome.

To generate the counterfactual weather scenarios, I begin with the status quo precipitation (Z^0) and create two types of patterns: 1) Mean Shift: $Z' = Z^0 \pm \zeta_1$, where $\zeta_1 \in [-0.25, 0.25]$. The generated Z' has the same variance as Z^0 but a different mean. 2) Variance Shift: $Sd(Z') = Sd(Z^0) \cdot \zeta_2$, where $\zeta_2 \in [0.75, 1.25]$. The generated Z' has the same median as Z^0 but a different variance. The latter is achieved by nonparametrically scaling the standard deviation from Z^0 while maintaining the median and clipping the result at 0.[27] I focus solely on changes in precipitation, assuming all other weather variables remain at their status quo levels. Throughout the analysis, I also hold household characteristics and income constant, with the exception of NDVI, which is updated through a reduced-form analysis as discussed in Section A.A4.[28] Using these counterfactuals, I employ the empirical model from Equation 12 to calculate the optimal price. Further details on the optimization procedure can be found in Section A.A4.

A. Welfare Results - Mean Shift (ζ_1)

This subsection presents the results of shifting the mean of precipitation by adding $\zeta_1 \in [-0.25, 0.25]$ to the status quo weather, showcasing the heterogeneous welfare effects across income strata, particularly when demand shifts rightward. Under this scenario, the utility expects the average precipitation to change by ζ_1 while the variance remains constant. The resulting menu of optimal prices can be found in Section A.A4.

Figure 9 shows the general welfare results for precipitation expected to be drier ($\zeta_1 \leq 0$) and rainier ($\zeta_1 \geq 0$). When $\zeta_1 \leq 0$, demand shifts rightward, the quantity constraint becomes binding, and the revenue constraint becomes slack. Even with price increases

[27] Although precipitation often follows a right-skewed log-normal distribution, a parametric assumption makes scaling by the factor ζ_2 imprecise. As noted by meteorologists Heredia et al. (2018), a nonparametric approach is more realistic for generating synthetic high-variance data.

[28] I also assume the income effect in the demand function (Equation 6) does not change with weather.

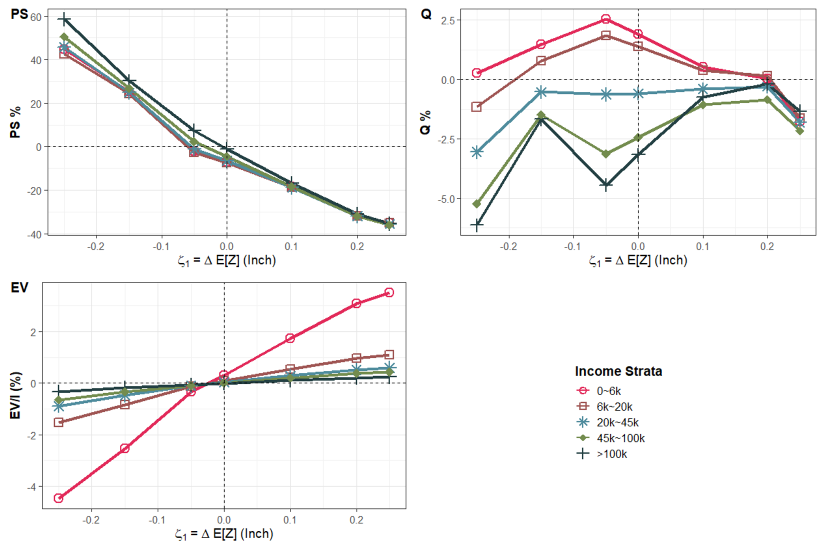

Figure 9. : Welfare from the Optimal Price in Shifting Mean ($\zeta_1 \in [-0.25, 0.25]$)

only for higher tiers (see Figure A7) and weighting the objective by EV/I, lower-income households exhibit a larger welfare loss compared to strata above \$20k, which show minimal loss. Conversely, when $\zeta_1 \geq 0$ (i.e., when the weather is rainier), demand shifts leftward, the revenue loss is more pressing, but due to the implementation of the loss function and lower weight ($\lambda = 0.5$), I still observe revenue losses. This results in a welfare increase for lower-income households.

One of the caveats from this analysis is that I don't directly observe the household-level income. Instead, the income is extrapolated through the house value and normalized by zip code, assuming the distributions of house value and income within each zip code are the same. To test the validity of this assumption, I limited the welfare result to 1) the top 50th percentile of house built year (built after 1982) and 2) houses built after 2000 ($\sim 27\%$ of the total households), and the result could be found in Section A.A4. To summarize, in both specifications, the lowest stratum still generates much lower welfare compared to other strata, and both results are at a similar level to the result from the full sample. This showcases that house value from the appraisal district is a good enough proxy for the household income.

B. Welfare Results - Variance Shift (ζ_2)

This subsection presents the results of shifting precipitation variance (ζ_2), demonstrating how changes in variance also shift the demand curve. Based on the result from Section III.D, lower variance ($\zeta_2 < 1$) will shift the demand curve to the left, while higher variance ($\zeta_2 > 1$) will shift the demand curve to the right. Under this scenario, the utility expects the variance of precipitation to change by a factor of ζ_2 while the median remains constant.

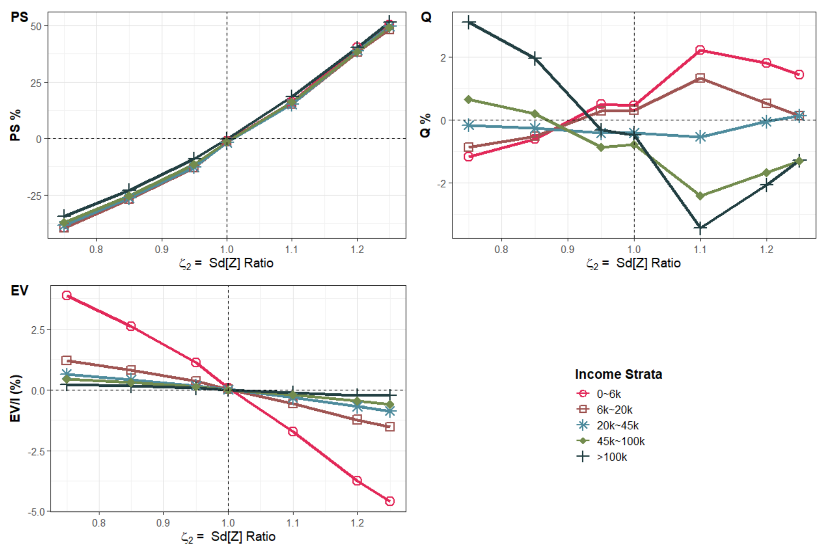

Figure 10. : Welfare from the Optimal Price in Shifting Variance ($\zeta_2 \in [0.75, 1.25]$)

Figure 10 showcases the general welfare results for $\zeta_2 \leq 1$ and $\zeta_2 \geq 1$. When $\zeta_2 \leq 1$ (variance decreases), demand shifts leftward, the revenue loss is more concerning, and the conservation constraint is slack, but again, due to the implementation of the loss function and $\lambda = 0.5$, I still observe revenue to be lower than the status quo when precipitation variance is very low. This occurs because, as variance decreases, the right-skewed precipitation distribution results in more rain during historically drier months. Consequently, welfare generally increases, particularly for the lowest-income stratum ($0 \sim \$6k$). Conversely, when $\zeta_2 > 1$ (variance increases), demand shifts rightward, the quantity constraint binds, and the revenue constraint is slack. The resulting different weather patterns decrease welfare for all strata, disproportionately affecting lower-income households (similar to the $\zeta_1 \leq 0$ case). The interplay of equity welfare weight (w_h), both constraints and weather, creates an interesting pattern for quantity. When precipitation variance is low, to reduce revenue loss, the utility will "prefer" more quantities (ceiling by the upper bound) to be consumed by higher income households. However, when precipitation variance is high, the pressure to maintain the conservation goal is largely achieved by the decreasing quantity for the higher-income households. As the equity weight is larger for the low-income households, they don't have the same "burden" as the high-income households to maintain the distributional goal.

Combining these results, a clear pattern emerges: the distributional effect of IBP is not improved, and is often worsened when weather shifts demand rightward, even with income-weighting in the objective function. Several factors explain this outcome: 1) Water bills naturally represent a higher proportion of income for lower-income households, so any price or weather change incurs a larger relative welfare shift (EV/I).[29] 2) The inverse elasticity rule suggests decreasing prices for high-elasticity consumers, which is distributionally counterproductive, as many high-elasticity households are in the highest income stratum and consume large quantities.[30] 3) The most important reason is that a non-negligible number of lower-income households are high-quantity consumers with relatively low elasticities, often placing them in higher price tiers than their income suggests.[31] As I explained above, the optimization process has tried its best to shift the burden of the conservation goal towards higher-income households, and we see a large decrease in quantity when demand shifts to the right. However, the process is only achieved through targeting quantity, instead of targeting income. The existence of high-quantity-low-income households will inevitably suffer from higher prices from the conservation constraint, which leads to lower welfare. Even though they are not the majority of the lowest income stratum, they still drag the results down.

All these reasons compromise the premise that income is a good indicator of quantity, and higher marginal prices in higher tiers promote equity. These effects work against the assumption of using income as a proxy for consumption in rate design, making IBP an unreliable tool for achieving distributional goals. This finding aligns with previous literature concluding that IBP changes often favor higher-income households (Ruijs, Zimmermann and van den Berg (2008), Echeverri (2023), Wichman (2024)).[32]

C. Shadow Cost of the Policy Constraints

The counterfactual analysis proves that, when weather shifts the demand curve right, it is difficult to improve or even maintain distributional equity while simultaneously achieving conservation goals. For the lowest-income stratum, this welfare imbalance stems from both the policy constraint and the extreme weather itself. While the weather is fully exogenous and the utility cannot control its welfare impact, policy constraints, on the other hand, can be mitigated by the utility. To isolate the welfare loss that can be mitigated by the utility, I first calculate the welfare impact of extreme weather alone by applying the status quo price ($\mathbf{p}_0$) to all counterfactual weather scenarios. The difference between that result and the welfare under the optimal price represents the welfare effect of the optimal pricing with their policy constraints. I perform this exercise specifically for the lowest-income stratum. The results for all scenarios can be found in Section A.A4.

For the mean-shift scenarios (ζ_1), when $\zeta_1 < 0$, the binding conservation constraint generates a shadow cost that reduces welfare. In the driest condition ($\zeta_1 = -0.25$), welfare under the status quo price is -2.41% on average. Compared to the optimal price result (-4.49%) from Figure 9, this implies the shadow cost of the conservation constraint incurs an average welfare loss of \$74.2 per household per month for the lowest

[29] Similar patterns have been pointed out by Ruijs, Zimmermann and van den Berg (2008) using data in Sao Paolo, Brazil.

[30] See Figures 4 and A5.

[31] See Figure A5.

[32] Both Ruijs, Zimmermann and van den Berg (2008) and Wichman (2024) have pointed out that IBP itself does not effectively target the lower stratum and that more active progressive measures are needed.

stratum. When $\zeta_1 > 0$, even though the revenue loss function is constructed as a policy constraint, the less strictness of following this constraint in reality made the optimal pricing and its distributional effect work as intended, creating a welfare increase from 2.18% to 3.52%, which is \$45.3 per household per month for the lowest stratum.

For the variance-shift scenarios (ζ_2), the effects are similar. In the lowest variance condition ($\zeta_2 = 0.75$), welfare under the status quo price is 2.45%. Compared to the optimal price result (3.88%) from Figure 10, this implies that a less strict revenue constraint enables the optimal pricing to create a welfare gain of \$48.5 per household per month for the lowest stratum. In the highest variance condition ($\zeta_2 = 1.25$), welfare under the status quo price is -2.48%. Compared to the optimal price result (-4.60%) from Figure 10, this implies a shadow cost from the conservation constraint of \$74.9 per household per month.

Both shifting mean and shifting variance showcase that the effectiveness of the optimal pricing procedure with accounting for welfare equity can be hindered by the strictness of the policy constraints. I modeled the policy behaviors of the utility using data and its goals. With the dwindling water supply, more metropolitan areas in the southwest regions of the US will be like Austin Water, adopting more and more strict conservation policies. Revenue, on the other hand, could be savaged by the city government as typically, the utility company is a public company. If the utility cares more about its revenue feasibility, i.e., λ is higher, a more strict revenue constraint could also impose a welfare burden for the lower-income households. In addition, I also check the robustness of the result with different segments of built years to ensure the estimated income from house appraisal value won't bias the result.

D. Zeroscaping for the Lowest Income Stratum

The preceding analysis concludes that under extreme weather, policy constraints on optimal prices cause greater welfare losses for the lowest-income stratum. Since these constraints are essential for utility operations, mitigating this welfare imbalance requires an additional policy targeted at this group. IBP's failure to achieve the distributional goal is largely due to the mismatch between income level and consumption level. To fully understand what variables could predict the consumption heterogeneity, I performed a reduced-form analysis (see Table A7), which shows that factors like house size, number of bathrooms, and vegetation health (NDVI) are significant predictors of high usage. Of these, NDVI becomes most pronounced during periods of low precipitation. Therefore, a natural policy to test is zeroscaping—reducing NDVI.

For the counterfactual, I halve the status quo NDVI for the lowest-income stratum, simulating the removal of roughly half their lawns,[33] and repeat the optimal pricing exercise. Since NDVI only changes for the lowest stratum, I present only their welfare changes.

Unsurprisingly, by halving NDVI, the lowest stratum consumes and pays less, resulting in an increase in welfare across all weather conditions, particularly when $\zeta_1 < 0$ and

[33]Note that I only modified NDVI > 0; if the status quo NDVI was already negative, I did not change it.

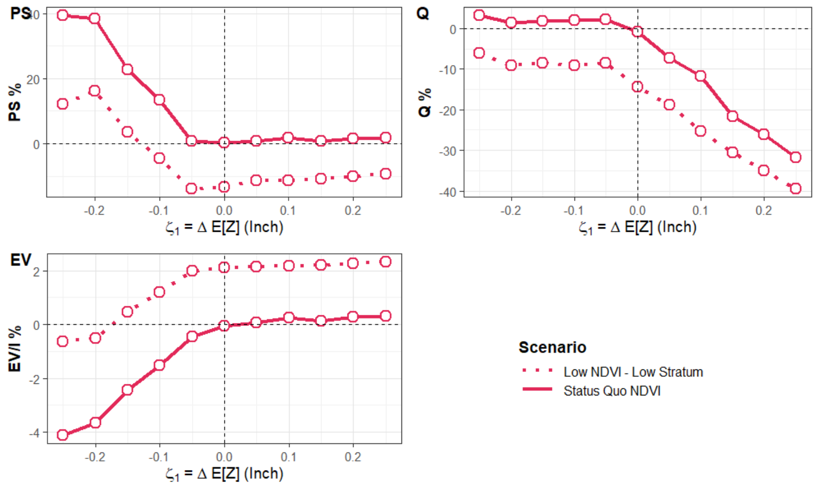

Figure 11. : Welfare Comparison of Halving NDVI ($\zeta_1 \in [-0.25, 0.25]$)

the demand curve shifts rightward. In the most extreme dry condition ($\zeta_1 = -0.25$), zeroscaping improves EV/I for the lowest stratum from -4.49% to -0.62%, corresponding to a gain of roughly \$121.58 per household per month. This measures the immediate welfare effect, not including long-term benefits.[34]

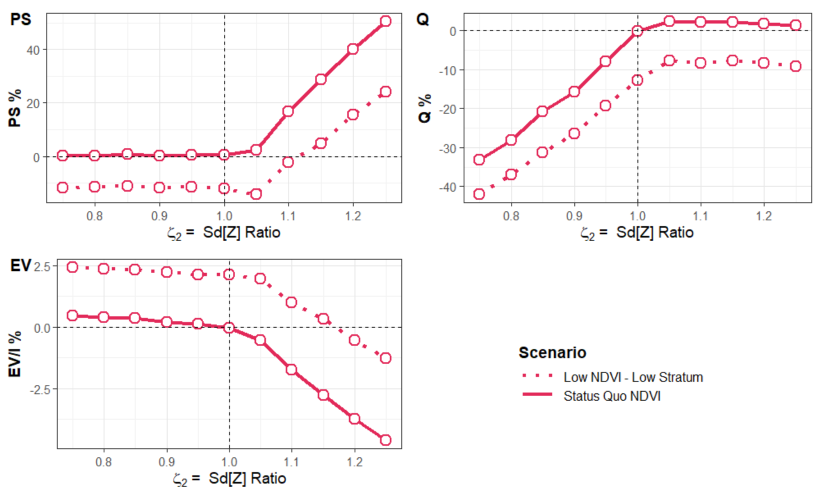

Figure 12. : Welfare Comparison of Halving NDVI ($\zeta_2 \in [0.75, 1.25]$)

[34]A one-time investment in zeroscaping would provide a near-permanent downward shift in the demand curve, increasing future welfare.

The pattern is similar for variance shifts. When demand shifts rightward ($\zeta_2 > 1$), xeriscaping improves welfare for the lowest stratum. In the extreme case of $\zeta_2 = 1.25$, it improves EV/I from -4.60% to -1.26%, corresponding to roughly \$116.46 per household per month.

To further investigate the value of this policy, I implemented multiple NDVI reduction levels, from 0 to 0.9 (representing a 90% lawn reduction). Figure 13 shows the welfare change for the two most extreme rightward-shift scenarios.

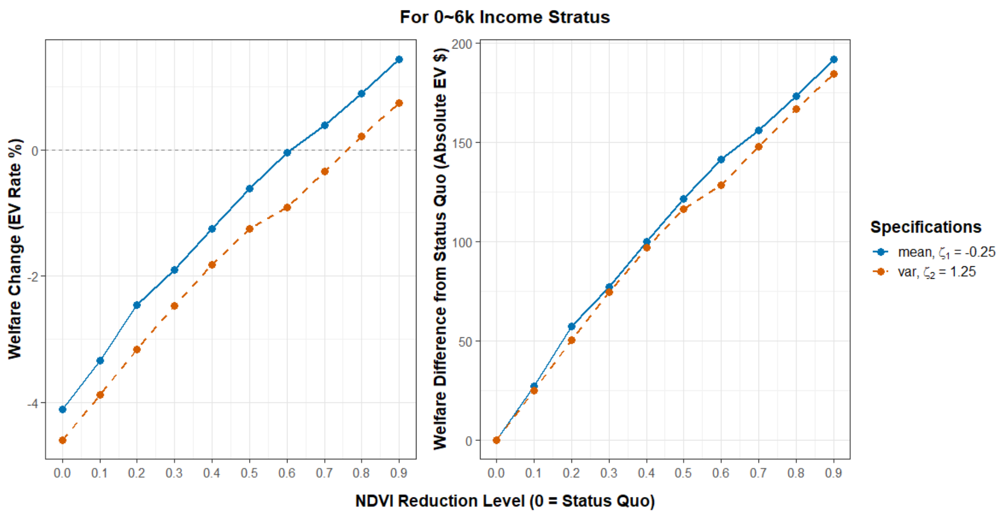

Figure 13. : Welfare Result under Extreme Weather Conditions by NDVI reduction level $\in [0, 0.9]$

The graph shows a clear increasing trend in welfare as the NDVI reduction level increases. For $\zeta_1 = -0.25$, a 0.3 reduction in NDVI generates \$75 per household per month, almost nullifying the welfare loss from the conservation constraint's shadow cost. A 0.6 reduction nullifies the welfare losses from both the shadow cost and the extreme weather itself. For $\zeta_2 = 1.25$, welfare is generally lower, meaning higher NDVI reductions are needed to achieve the same welfare level. On average, across all reduction levels, the $\zeta_2 = 1.25$ scenario generates \$6.06 less welfare per household per month, showcasing that higher variance represents a more severe extreme weather condition. Even so, zeroscaping remains a viable policy solution to improve the distributional effect.

VI. Conclusion

In this paper, I estimate the optimal Inclining Block Price (IBP) tariffs for water utilities using a Ramsey-style pricing model that incorporates extreme precipitation patterns. My initial reduced-form analysis reveals that deviations in precipitation are correlated with greater volatility in water consumption for higher-income households. This evidence highlights a significant price-setting challenge for utilities, particularly concerning their distributional goals.

I then employ a structural Discrete-Continuous Choice (DCC) model, applied to the piecewise-linear budget constraint, which incorporates a satellite-derived vegetation health index (NDVI) for household lawns. This allows for a more granular understanding of outdoor water usage in single-family households—a substantial portion of consumption that is most sensitive to precipitation.

Subsequently, I develop an empirical Ramsey model that includes both revenue recovery and quantity conservation constraints to determine the optimal IBP structure. Given that prices are set well before weather events occur, I explore optimal pricing under various weather scenarios, including shifts in the mean and variance of precipitation from the status quo. The analysis shows that when precipitation averages decrease or when volatility increases, the demand curve shifts rightward. This places more pressure on the now-binding conservation constraint, compounding the welfare loss caused by the extreme weather itself. Notably, the lowest-income stratum experiences the highest welfare losses. My findings confirm existing literature that current IBP tariffs used by utilities generally favor high-income households due to mismatches between household income and consumption levels. Furthermore, high-income households demonstrate greater price sensitivity during periods of low precipitation, further undermining the intended distributional effects of IBP.

This paper provides empirical measurements of the shadow cost of policy constraints in the utility's optimal rate design by comparing welfare outcomes with and without the conservation and revenue requirements. I separate the welfare loss imposed by the utility through its policies and the welfare loss from the weather, which is not controlled by the utility. I find that the lowest-income stratum experiences the largest welfare loss from these shadow costs, particularly when weather shifts the demand curve rightward. During these weather conditions, such as extreme drought or high precipitation variance, the conservation constraint becomes binding. This imposes an average welfare loss of \$74.2 and \$74.9 per household per month, respectively, on the lowest-income stratum. Notably, the shadow costs of these constraints do not cause significant welfare differences for other income strata. This finding underscores that price alone is an insufficient instrument for achieving a utility's multifaceted policy goals, as necessary conservation goal weaken the intended distributional effect.

This result has a profound policy implication: when weather cause a rising demand, using price as the sole instrument to achieve a utility's policy goals is not feasible. Additional policies are necessary to achieve distributional objectives. I focus on zeroscaping/xeriscaping (i.e., reducing lawn water needs and thus NDVI) as a straightforward policy to curb demand when the demand curve is pushed rightward. Under extremely dry conditions, reducing NDVI by 0.3 could generate \$75 in welfare per household per month, almost nullifying the welfare loss from the shadow cost of the conservation constraint. While high precipitation variance causes a larger welfare reduction overall, making the effect of zeroscaping slightly smaller, it remains highly effective.

Looking ahead, I plan to expand this research in two directions. First, I will model consumption smoothing programs. The existence of "Budget Billing" and similar programs reflects consumer risk aversion and a preference for stable monthly payments, particularly among lower-income individuals. This preference acts as a form of insurance against weather stochasticity and interacts significantly with the utility's pricing optimization problem. Second, I will endogenize vegetation changes (NDVI). I currently treat lagged NDVI and any reduction from zeroscaping as exogenous variables. However, consumers make long-term decisions about landscaping based on potential welfare gains. Endogenizing this decision-making process would allow for a more precise estimation of the long-term welfare benefits of xeriscaping.

REFERENCES

Atkinson, Anthony B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory*, 2(3): 244–263.

Baumgartner, Simon, Alex Stomper, Thomas Schober, and Rudolf Winter-Ebmer. 2022. "Banking on snow: Bank capital, risk, and employment."

Borenstein, Severin. 2009. "To what electricity price do consumers respond? Residential demand elasticity under increasing-block pricing." *Preliminary Draft April*, 30: 95.

Borenstein, Severin, and James B Bushnell. 2022. "Do two electricity pricing wrongs make a right? Cost recovery, externalities, and efficiency." *American Economic Journal: Economic Policy*, 14(4): 80–110.

Burtless, Gary, and Jerry A Hausman. 1978. "The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment." *Journal of political Economy*, 86(6): 1103–1130.

Castro-Rodríguez, Fidel, José María Da-Rocha, and Pedro Delicado. 2002. "Desperately Seeking θ's: Estimating the Distribution of Consumers under Increasing Block Rates." *Journal of Regulatory Economics*, 22(1): 29–58.

Coady, David, and Jean Drèze. 2002. "Commodity taxation and social welfare: The generalized Ramsey rule." *International Tax and Public Finance*, 9: 295–316.

Dubin, Jeffrey A, and Daniel L McFadden. 1984. "An econometric analysis of residential electric appliance holdings and consumption." *Econometrica: Journal of the Econometric Society*, 345–362.

Echeverri, Manuel Felipe Rojas. 2023. "Distributional Effects of a Nonlinear Price Scheme in Public Utilities."

Feldstein, Martin S. 1972. "Distributional Equity and the Optimal Structure of Public Prices." *The American Economic Review*, 62(1/2): 32–38.

Goulder, Lawrence H, Marc AC Hafstead, GyuRim Kim, and Xianling Long. 2019. "Impacts of a carbon tax across US household income groups: What are the equity-efficiency trade-offs?" *Journal of Public Economics*, 175: 44–64.

Gramazio, Connor C., David H. Laidlaw, and Karen B. Schloss. 2017. "Colorgorical: creating discriminable and preferable color palettes for information visualization." *IEEE Transactions on Visualization and Computer Graphics*.

Hall, BH, RE Hall, and JA Hausman. 1974. "Estimation and inference in nonlinear structural models." *Annals of economic and social measurement*, 3: 653–666.

Hanemann, W Michael. 1984. "Discrete/continuous models of consumer demand." *Econometrica: Journal of the Econometric Society*, 541–561.

Hausman, Jerry A. 1979. "The effect of wages, taxes, and fixed costs of women's labor force participation."

Hausman, Jerry A. 1981. "Exact consumer's surplus and deadweight loss." *The American Economic Review*, 71(4): 662–676.

Heredia, María Belén, Clémentine Junquas, Clémentine Prieur, and Thomas Condom. 2018. "New statistical methods for precipitation bias correction applied to WRF model simulations in the Antisana region, Ecuador." *Journal of Hydrometeorology*, 19(12): 2021–2040.

Hewitt, Julie A, and W Michael Hanemann. 1995. "A discrete/continuous choice approach to residential water demand under block rate pricing." *Land Economics*, 173–192.

Holland, Stephen P, and Erin T Mansur. 2008. "Is real-time pricing green? The environmental impacts of electricity demand variance." *The Review of Economics and Statistics*, 90(3): 550–561.

Ito, Koichiro. 2014. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing." *American Economic Review*, 104(2): 537–563.

Känzig, Diego R. 2023. "The unequal economic consequences of carbon pricing." National Bureau of Economic Research.

Kim, Kyungmin. 2020. "Income inequality and house prices in the United States: A panel VAR analysis." *Economics Bulletin*, 40(3): 2111–2120.

Lin, Chen, Thomas Schmid, and Michael S Weisbach. 2017. "Price risk, production flexibility, and liquidity management: Evidence from electricity generating firms." National Bureau of Economic Research.

Nataraj, Shanthi, and W Michael Hanemann. 2011. "Does marginal price matter? A regression discontinuity approach to estimating water demand." *Journal of Environmental Economics and Management*, 61(2): 198–212.

Nielsen-Gammon, John, Jacob Escobedo, Catherine Ott, Jeramy Dedrick, and Ali Van Fleet. 2020. "Assessment of historic and future trends of extreme weather in Texas, 1900-2036."

Olmstead, Sheila M. 2009. "Reduced-form versus structural models of water demand under nonlinear prices." *Journal of Business & Economic Statistics*, 27(1): 84–94.

Olmstead, Sheila M, W Michael Hanemann, and Robert N Stavins. 2007. "Water demand under alternative price structures." *Journal of Environmental economics and management*, 54(2): 181–198.

Parry, Ian WH, and Roberton C Williams III. 2010. "What are the costs of meeting distributional objectives for climate policy?" *The BE Journal of Economic Analysis & Policy*, 10(2).

Pint, Ellen M. 1999. "Household responses to increased water rates during the California drought." *Land economics*, 246–266.

Ragonneau, T. M. 2022. "Model-Based Derivative-Free Optimization Methods and Software." PhD diss. Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China.

Ragonneau, T. M., and Z. Zhang. 2025. "COBYQA Version 1.1.2."

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *The economic journal*, 37(145): 47–61.

Reiss, Peter C, and Matthew W White. 2005. "Household electricity demand, revisited." *The Review of Economic Studies*, 72(3): 853–883.

Ruijs, Arjan. 2009. "Welfare and distribution effects of water pricing policies." *Environmental and Resource Economics*, 43: 161–182.

Ruijs, Arjan, Anja Zimmermann, and Marrit van den Berg. 2008. "Demand and distributional effects of water pricing policies." *Ecological Economics*, 66(2-3): 506–516.

Strand, Jon, and Ian Walker. 2005. "Water markets and demand in Central American cities." *Environment and Development Economics*, 10(3): 313–335.

Szabo, Andrea. 2015. "The value of free water: analyzing South Africa's free basic water policy." *Econometrica*, 83(5): 1913–1961.

Train, Kenneth, Patrice Ignelzi, Robert Engle, Clive Granger, and Ramu Ramanathan. 1984. "The billing cycle and weather variables in models of electricity sales." *Energy*, 9(11-12): 1041–1047.

Wichman, Casey J. 2024. "Efficiency, Equity, and Cost-Recovery Trade-Offs in Municipal Water Pricing." Resources for the Future Working Paper DP 24-18.

Wolak, Frank A. 2016. “Designing nonlinear price schedules for urban water utilities to balance revenue and conservation goals.” National Bureau of Economic Research.

Zhang, Fudong. 2016. “Inequality and House Prices.” *Journal of Political Economy (forthcoming/working paper status at the time).*

APPENDIX

All plots of this paper have utilized the software "Colorgorical" to choose a more clear and aesthetically pleasing color scheme. Gramazio, Laidlaw and Schloss (2017). All codes producing the results and graphs related to this project are posted here.

A1. Water Utility Pricing in Austin, TX

PRECIPITATION TREND IN AUSTIN

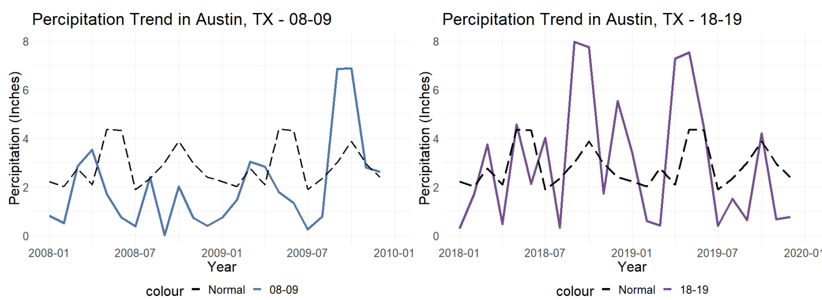

Figure A1. : Precipitation Trends of 2008-2009 vs. 2018-2019

CURRENT PRICING STRUCTURE

Table A1—: Pricing Structure

Kink Points (kGallon)	Fixed Charge ($)	Marginal Price ($)	
		Non-CAP	CAP
0-2	8.5	3.09	2.42
2-6	10.8	5.01	4.1
6-11	16.5	8.54	6.72
11-20	37	12.9	11.56
>20	37	14.41	14.26

Note: Each fixed charge is composed of fixed payment and meter charge. The fixed payment depends on final consumption quantity, and the meter charge assumes a 5/8 meter size, which is the most common residential meter size. The marginal price includes volume charge, reserved fund charge and community benefit charge. These itemized charges are all billed per 1,000 gallons, meaning they are essentially part of the marginal price. Volume charge depends on the amount of final consumption quantity while the other two are billed per 1,000 gallons. Reserved fund surcharge goes into a restricted reserve fund to offset water service revenue shortfalls that may impact operations and services. Community benefit charge is only billed to Non-CAP consumers to fund the CAP.

Tier and Income Distribution

Based on the status quo price and status quo quantity distribution, I divide the household's monthly income into 5 different strata. Since both the quantity and income are very right-skewed, only for this plot, I filtered out all households with monthly household income bigger than $\$250,000$, and quantities higher than $100,000$ gallons. These households are not filtered out in demand estimation and counterfactual analysis, but are simply filtered out for the picture below for visualization purposes.

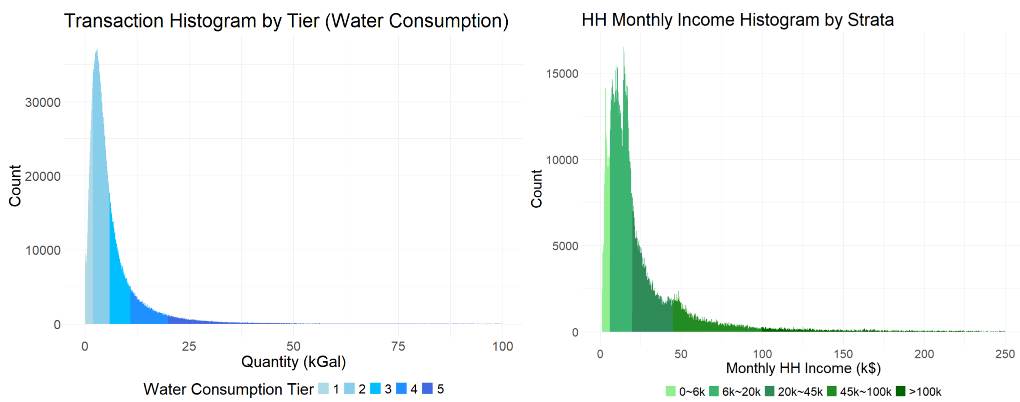

Figure A2. : Monthly Water Quantity Distribution vs. Income Distribution

Table A2—: Monthly Water Quantity Distribution vs. Income Distribution

Income Strata ($)	Percentage	Tiers (kGal)	Percentage
$0 \sim 6k$	0.146	$0 \sim 2k$	0.159
$6k \sim 20k$	0.489	$2k \sim 6k$	0.503
$20k \sim 45k$	0.216	$6k \sim 11k$	0.200
$45k \sim 100k$	0.104	$11k \sim 20k$	0.0927
$> 100k$	0.0441	$> 20k$	0.0456

Note: Just like the utility targeting different quantity levels of households. This division of household monthly income into different strata is without loss of generality and serves as a way to gauge the distributional effect incurred by any price change.

Descriptive and Reduced Form Evidence

In the interest of seeing how different income strata will react to abnormal weather changes and the descriptive evidence of Figure 2, I performed the following OLS:

$$\Delta q = \beta_0 + \beta_1 \Delta Z_T + \beta_2 \Delta Z_P + \beta_3 \text{Income} * \Delta Z_T + \beta_4 \text{Income} * \Delta Z_P + \alpha p + \varepsilon$$

where Income is the income strata defined in Table A2. ΔZ for both precipitation and temperature is the difference between observed Z_m for month m and the corresponding 30 year average. $\Delta q = \frac{q_m - \bar{q_m}}{\bar{q_m}}$, which is the quantity deviation from the average quantity for the specific household for a specific month between 2016-2020 in percentage terms.

Table A3—: Regression Results for Quantity Deviation

Variable	Estimate	Std. Error	t value	Pr(> $\|t\|$)
(Intercept)	−0.4182	0.0008490	−492.609	< 2e-16 ***
ΔZ_P	−0.01668	0.0003582	−46.555	< 2e-16 ***
ΔZ_T	0.004787	0.0001438	33.283	< 2e-16 ***
income_strata: 6k∼20k	−0.02104	0.0007867	−26.739	< 2e-16 ***
income_strata: 20k∼45k	−0.08237	0.0008990	−91.627	< 2e-16 ***
income_strata: 45k∼100k	−0.1673	0.001096	−152.710	< 2e-16 ***
income_strata: >100k	−0.2476	0.001513	−163.593	< 2e-16 ***
p	0.07486	0.00008868	844.089	< 2e-16 ***
ΔZ_P×income_strata: 6k∼20k	0.001576	0.0004098	3.846	0.00012 ***
ΔZ_P×income_strata: 20k∼45k	−0.009682	0.0004614	−20.986	< 2e-16 ***
ΔZ_P×income_strata: 45k∼100k	−0.01923	0.0005504	−34.933	< 2e-16 ***
ΔZ_P×income_strata: >100k	−0.01874	0.0007744	−24.198	< 2e-16 ***
ΔZ_T×income_strata: 6k∼20k	0.00009410	0.0001657	0.568	0.57012
ΔZ_T×income_strata: 20k∼45k	0.0004584	0.0001880	2.439	0.01474 *
ΔZ_T×income_strata: 45k∼100k	−0.002830	0.0002234	−12.667	< 2e-16 ***
ΔZ_T×income_strata: >100k	−0.002834	0.0003005	−9.428	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.4014 on 2351610 degrees of freedom
Multiple R-squared: 0.263, Adjusted R-squared: 0.263
F-statistic: 5.596e+04 on 15 and 2351610 DF, p-value: < 2.2e-16

In this model, controlling for differences in precipitation, maximum temperature, and the marginal price, the relationship between income strata and the percentage difference in water consumption is seen through: 1) for baseline differences (when Δ weather and price are zero), compared to the lowest income strata (0-6k), most higher income strata show a statistically significant lower baseline percentage difference in water consumption from their usual amount. This does not mean higher income groups use less water overall; it means their deviation from their own usual quantity, under these specific baseline conditions, is lower than the deviation for the reference group. This effect is likely influenced by the strong role of price in this model, as marginal price might be correlated with income or consumption levels that influence which price tier is reached. 2) The linear effect of precipitation difference: the interaction terms between income strata and ΔZ_P are largely significant. This means that the linear rate at which Δq changes for every unit increase in precipitation difference varies significantly across income strata. In particular, higher income strata (20k-45k, 45k-100k, >100k) show increasingly more negative interaction terms (-0.009682, -0.01923, -0.01874). This means the negative impact of the additional precipitation difference on consumption deviation is increasingly stronger as income rises in these groups. 3) the linear effect of temperature difference: many interaction terms between income strata and ΔZ_T are also significant, indicating that the linear rate at which Δq changes for every unit increase in temperature difference varies significantly across income strata. However, the relationship is less clear-cut compared to the one from precipitation. Compared to the reference group (where the effect of ΔZ_T is 0.004787, meaning a 0.48 percentage point increase in consumption difference for every unit increase in temperature difference), the 6k-20k stratum's interaction is not significant (0.0000941), suggesting its linear temperature sensitivity is not statistically different from the reference group in this model. The 20k-45k stratum has a significant positive interaction (0.0004584), making its positive temperature effect slightly stronger (0.004787 + 0.0004584 = 0.0052454). The 45k-100k and >100k strata have significant negative interaction terms (-0.002830 and -0.002834). This means that although the base effect of temperature difference is positive, the additional positive effect seen in higher income groups in the first model is now appearing as a reduction in the sensitivity compared to the reference group's base sensitivity (0.004787 - 0.002830 = 0.001957 for 45k-100k; 0.004787 - 0.002834 = 0.001953 for >100k).

This difference between temperature and precipitation prompts the research to be more focused on changing precipitation. In addition, a structural model is used to estimate the demand on the panel data to show how price, precipitation, and the interaction term with income strata will affect the demand.

A2. Demand Model and Estimation

Additional Case in the Unconditional Indirect Utility Function

When constructing the unconditional indirect utility function from the conditional ones, the case where consumption occurs exactly at a tier boundary (kink point), rather than strictly within a tier, may require clarification. Continuing with the two-tier scenario ($K = 2$) without loss of generality, this case arises when the household's optimal consumption calculated using the tier 1 price, $g(p_1, I)$, would exceed the tier 1 limit q_1, *and* the optimal consumption calculated using the tier 2 price and virtual income, $g(p_2, I + d_2)$, would fall below q_1. That is, the condition is $g(p_2, I + d_2) \leq q_1 < g(p_1, I)$ (assuming $d_1 = 0$ or defined appropriately). To satisfy the Incentive Compatibility (IC) constraint (i.e., ensure the chosen consumption is utility-maximizing given the full budget set), the household optimally consumes exactly at the kink point q_1. The diagram below illustrates this:

As shown, under the condition $g(p_2, I + d_2) \leq q_1 < g(p_1, I)$, the optimal (incentive compatible) choice for this household is to consume the bundle corresponding to the kink point, $(q_1, I - \text{Bill}(q_1))$. Here, $\text{Bill}(q_1)$ represents the total water bill incurred when consuming exactly q_1 units (e.g., typically $A_2 + p_1 q_1$ in a two-tier system where A_2 is the fixed charge associated with entering tier 2). If consumption at the kink were not allowed as an option in the model, such consumers would be incorrectly assigned to

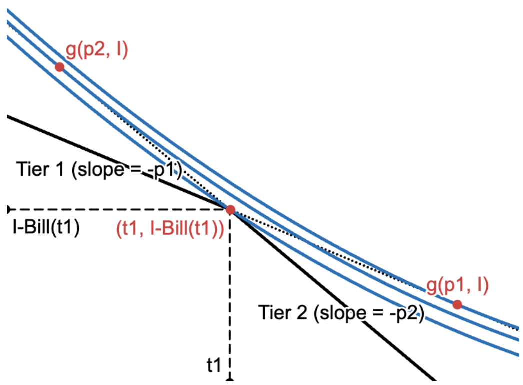

Figure A3. : The household consumes exactly at the kink $(q_1, I - \text{Bill}(q_1))$

consume within one of the adjacent tiers, violating the true optimum.

Generalizing, for any tier boundary q_k where $k \in \{1, \ldots, K-1\}$, some households might satisfy the condition $g(p_{k+1}, I + d_{k+1}) \leq q_k < g(p_k, I + d_k)$, for whom the optimal consumption point is exactly $(q_k, I - \text{Bill}(q_k))$. In our empirical application to Austin, which has 5 tiers ($K = 5$), there are 4 such kink points (q_1, q_2, q_3, q_4), requiring the model to allow for consumption exactly at these quantities to ensure incentive compatibility across the full range of consumption.

Likelihood Function

Without loss of generality, I explain the derivation of the likelihood function for a two-tier case ($K = 2$); the approach generalizes to K tiers. Conditional on the ex-ante optimal choice (based on the known preference shock η but ignoring the ex-post error ν) being tier 1 ($k^* = 1$), the condition is $\log w_1 + \eta \leq \log q_1$. Similarly, $k^* = 2$ if $\log w_2 + \eta > \log q_1$. The remaining possibility, ensuring incentive compatibility and covering all η, is consumption at the kink q_1, which occurs if $\log w_2 + \eta \leq \log q_1 < \log w_1 + \eta$. Rearranging these conditions on η and adding the ex-post error ν (unobserved by the household when choosing the tier/kink), the observed log-consumption $\log(w)$ is:

$$\log(w) = \begin{cases} \log(w_1) + \eta + \nu & \text{if } \eta \leq \log(q_1) - \log(w_1) \\ \log(q_1) + \nu & \text{if } \log(q_1) - \log(w_1) < \eta \leq \log(q_1) - \log(w_2) \\ \log(w_2) + \eta + \nu & \text{if } \eta > \log(q_1) - \log(w_2) \end{cases}$$

which is the two-tier special case of Equation 7. Note that since the household does not observe ν when making its choice, ν affects the final observed consumption in all scenarios.

Let $f(\cdot)$ denote a probability density function (pdf). The overall likelihood for an observation $\log w$ is the sum of the contributions from these three mutually exclusive and exhaustive scenarios ($L = L_1 + L_{kink} + L_2$).

Case 1: $k^* = 1$ ($\eta \leq \log(q_1) - \log(w_1)$) The observed log consumption is $\log w = \log w_1 + \eta + \nu$. The contribution to the likelihood depends on the joint pdf of $(\eta + \nu, \eta)$, integrated over the relevant range of η:

$$L_1 = \int_{-\infty}^{\log q_1 - \log w_1} f_{\nu+\eta,\eta}(\log w - \log w_1, \eta)\, d\eta$$

Case 2: $k^* = 2$ ($\eta > \log(q_1) - \log(w_2)$) The observed log consumption is $\log w = \log w_2 + \eta + \nu$. The likelihood contribution is:

$$L_2 = \int_{\log q_1 - \log w_2}^{\infty} f_{\nu+\eta,\eta}(\log w - \log w_2, \eta)\, d\eta$$

Case 3: Kink Consumption ($\log(q_1) - \log(w_1) < \eta \leq \log(q_1) - \log(w_2)$) The observed log consumption is $\log w = \log q_1 + \nu$. The likelihood contribution depends on the joint pdf of (ν, η), integrated over the relevant range of η:

$$L_{kink} = \int_{\log q_1 - \log w_1}^{\log q_1 - \log w_2} f_{\nu,\eta}(\log w - \log q_1, \eta)\, d\eta$$

If I assume $\eta \sim N(0, \sigma_\eta^2)$ and $\nu \sim N(0, \sigma_\nu^2)$, and that they are independent, these integrals can be solved in closed form. Let $\phi(\cdot)$ and $\Phi(\cdot)$ be the standard normal pdf and cdf, respectively. The joint distribution of $(\eta, \nu + \eta)$ is bivariate normal. Let $\rho_s = \text{Corr}(\eta, \nu + \eta) = \sigma_\eta / \sqrt{\sigma_\eta^2 + \sigma_\nu^2}$.

Evaluating the first integral (L_1):

$$\begin{aligned}
L_1 &= \int_{-\infty}^{\log q_1 - \log w_1} f_{\nu+\eta}(\log w - \log w_1) f_{\eta|\nu+\eta}(\eta | \log w - \log w_1)\, d\eta \\
&= f_{\nu+\eta}(\log w - \log w_1) \int_{-\infty}^{\log q_1 - \log w_1} f_{\eta|\nu+\eta}(\eta | \log w - \log w_1)\, d\eta \\
&= \frac{1}{\sqrt{\sigma_\eta^2 + \sigma_\nu^2}} \phi\left(\frac{\log w - \log w_1}{\sqrt{\sigma_\eta^2 + \sigma_\nu^2}}\right) \Phi\left(\frac{(\log q_1 - \log w_1)/\sigma_\eta - \rho_s(\frac{\log w - \log w_1}{\sqrt{\sigma_\eta^2 + \sigma_\nu^2}})}{\sqrt{1 - \rho_s^2}}\right) \\
&\equiv \frac{\phi(s_1)}{\sqrt{\sigma_\eta^2 + \sigma_\nu^2}} \Phi(r_1)
\end{aligned}$$

where $s_1 = (\log w - \log w_1)/\sqrt{\sigma_\eta^2 + \sigma_\nu^2}$ and $r_1 = (t_1^* - \rho_s s_1)/\sqrt{1 - \rho_s^2}$ with $t_1^* = (\log q_1 - \log w_1)/\sigma_\eta$.

Similarly, evaluating the integral for the second case (L_2):

$$
\begin{aligned}
L_2 &= f_{\nu+\eta}(\log w - \log w_2) \int_{\log q_1 - \log w_2}^{\infty} f_{\eta|\nu+\eta}(\eta | \log w - \log w_2)\, d\eta \\
&= f_{\nu+\eta}(\log w - \log w_2) \left[1 - \Phi\left(\frac{(\log q_1 - \log w_2)/\sigma_\eta - \rho_s\left(\frac{\log w - \log w_2}{\sqrt{\sigma_\eta^2+\sigma_\nu^2}}\right)}{\sqrt{1-\rho_s^2}} \right) \right] \\
&\equiv \frac{\phi(s_2)}{\sqrt{\sigma_\eta^2+\sigma_\nu^2}} (1 - \Phi(n_2))
\end{aligned}
$$

where $s_2 = (\log w - \log w_2)/\sqrt{\sigma_\eta^2+\sigma_\nu^2}$ and $n_2 = (m_1 - \rho_s s_2)/\sqrt{1-\rho_s^2}$ with $m_1 = (\log q_1 - \log w_2)/\sigma_\eta$.

Finally, evaluating the integral for the third case (kink consumption, L_{kink}), using the independence of ν and η:

$$
\begin{aligned}
L_{kink} &= \int_{t_1^*\sigma_\eta}^{m_1\sigma_\eta} f_\nu(\log w - \log q_1) f_\eta(\eta)\, d\eta = f_\nu(\log w - \log q_1) \int_{t_1^*\sigma_\eta}^{m_1\sigma_\eta} f_\eta(\eta)\, d\eta \\
&= \frac{1}{\sigma_\nu} \phi\left(\frac{\log w - \log q_1}{\sigma_\nu} \right) \left[\Phi\left(\frac{m_1\sigma_\eta}{\sigma_\eta} \right) - \Phi\left(\frac{t_1^*\sigma_\eta}{\sigma_\eta} \right) \right] \\
&= \frac{\phi(u_1)}{\sigma_\nu} (\Phi(m_1) - \Phi(t_1^*))
\end{aligned}
$$

where $u_1 = (\log w - \log q_1)/\sigma_\nu$.

Summing the three components gives the likelihood for the two-tier case:

$$
\begin{aligned}
L &= L_1 + L_{kink} + L_2 \\
&= \frac{\phi(s_1)}{\sqrt{\sigma_\eta^2+\sigma_\nu^2}} \Phi(r_1) + \frac{\phi(u_1)}{\sigma_\nu} (\Phi(m_1) - \Phi(t_1^*)) + \frac{\phi(s_2)}{\sqrt{\sigma_\eta^2+\sigma_\nu^2}} (1 - \Phi(n_2))
\end{aligned}
$$

Using the general definitions $t_k^* = (\log q_k - \log w_k)/\sigma_\eta$, $m_k = (\log q_k - \log w_{k+1})/\sigma_\eta$, $s_k = (\log w - \log w_k)/\sqrt{\sigma_\eta^2+\sigma_\nu^2}$, $u_k = (\log w - \log q_k)/\sigma_\nu$, $r_k = (t_k^* - \rho_s s_k)/\sqrt{1-\rho_s^2}$, $n_k = (m_{k-1} - \rho_s s_k)/\sqrt{1-\rho_s^2}$, and applying boundary conditions ($m_0 \to -\infty \implies n_1 \to -\infty \implies \Phi(n_1) = 0$; $q_2 \to \infty \implies t_2^* \to \infty \implies r_2 \to \infty \implies \Phi(r_2) = 1$), this derived likelihood function matches the general form given in Equation 8 for $K = 2$.

NDVI

The use of the Normalized Difference Vegetation Index (NDVI) for estimating water demand was introduced by Wolak (2016). The essential idea is to quantify vegetation health for specific geographical locations – in this case, the health of lawns associated with single-family homes. Higher NDVI values indicate healthier vegetation, suggesting greater lawn care efforts.

NDVI is calculated from satellite imagery using the difference between the reflectance in the near-infrared (NIR) and red light bands. Healthy plants typically reflect more NIR light and absorb more red light (appearing green), while less healthy or stressed plants reflect less NIR and more red light (appearing yellow or brown). The index is typically calculated as (NIR - Red) / (NIR + Red), normalized to a range between -1 and +1, where higher values indicate healthier, denser vegetation. Values near -1 often correspond to water bodies. Values around 0 typically represent bare soil or sparse vegetation, and values approaching +1 indicate dense, healthy vegetation like forests or well-maintained lawns.

The raw data I use is the Sentinel-2 Surface Reflectance dataset,[35] providing imagery with a 10m x 10m spatial resolution (pixel size). Since typical residential lots in Texas are larger than this pixel size, this higher resolution allows for a more precise calculation of NDVI within each lot compared to the 30m x 30m resolution used by Wolak (2016). Sentinel-2 images are collected frequently (on average, every 5 days for a given location), though not always regularly, resulting in multiple images per month. Some images, however, contain areas obscured by clouds. To address this, I employ cloud masking and create monthly composite images.[36] The cloud masking algorithm utilizes bands of the images, such as the "Cloud Probability" or "QA". These bands indicate the likelihood of cloud cover or contain bitwise flags for clouds and shadows. To create a cloud-free composite, I apply these masks to filter out cloudy pixels from multiple images over all images of a calendar month of the same area, then merge the remaining clear-sky pixels into a seamless, cloud-free composite.

To visually illustrate the process of calculating NDVI in Figure A4, I take a small sample from the Brentwood neighborhood in Austin, near the intersection of W Koenig Lane and Burnet Road. While detecting subtle vegetation differences can be difficult in the raw satellite image, the corresponding NDVI visualization clearly showcases the variations in vegetation health between households.

Summary Statistics

After matching the panel data with TCAD records, filtering outliers, and removing households located outside Travis County (and thus likely outside the Austin Water service area), the final dataset primarily covers the period from May 2018 to December 2019. It is important to note that this data includes households on slightly different payment plans. Specifically, some households meeting certain income requirements are eligible for CAP, which offers lower marginal prices. The inclusion of CAP households provides valuable price variation for model identification, despite their small number. The summary statistics can be found in Table A4.[37]

[35] If using Google Earth Engine to access data, please refer to Google Earth Engine Sentinel-2 Surface Reflectance.

[36] Details for cloud masking in Google Earth Engine can be found here Cloud Masking - gee map.

[37] Overall, approximately 60% of households and 55.84% of transactions from the original data are used for the demand estimation analysis. Among the final sample, 99.57% are non-CAP consumers, and 0.43% are CAP consumers.

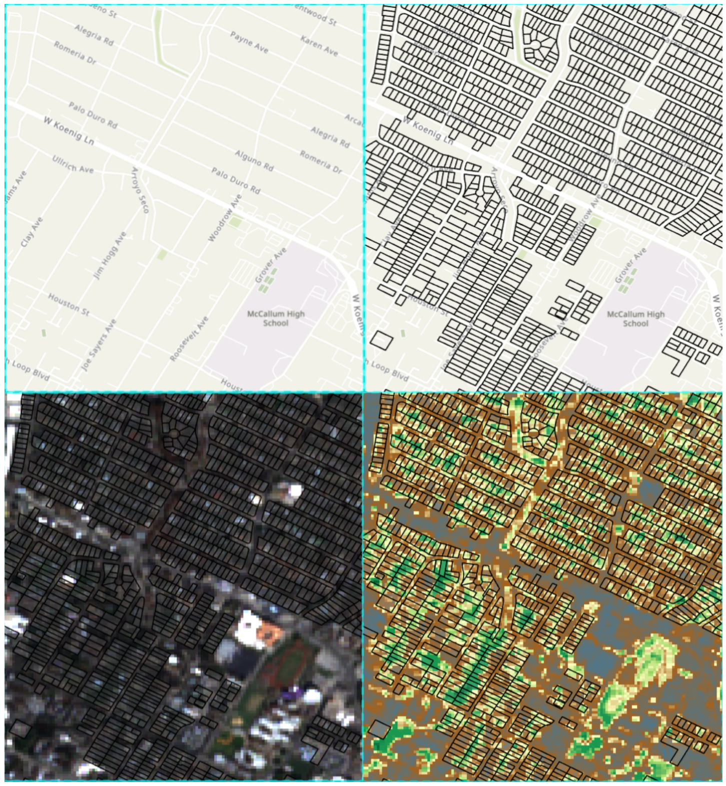

Figure A4. : The Process of Calculating NDVI

Note: From Top Left to Bottom Right: OpenStreetMap base map; OpenStreetMap with household lot shapefiles overlaid; Sentinel-2 Raw Image mosaic with shapefiles overlaid; and final NDVI visualization derived from Sentinel-2 with shapefiles overlaid.

Table A4—: Summary Statistics

Parameter Name	Min	1st Quartile	Median	Mean	3rd Quartile	Max	N
Heavy Water Appliances	0	0	0	0.1637	0	12	127320
Bedrooms	1	1	1	1.439	1	31	127320
Bathrooms	0.1	2	2	2.427	2.5	34	127320
Lot Size (Acre)	0.02136	0.15018	0.19060	0.27396	0.25309	278.25085	127320
Household Monthly Income (k$)	1.2	8.642	15.320	29.295	27.765	5057.051	127320
NDVI	-0.3564	0.3266	0.4038	0.3988	0.4766	0.7729	2351626
Mean Max Temp (F)	56.24	68.98	83.74	82.18	95.04	101.65	2351626
IQR Max Temp (F)	2	4.984	9.499	9.904	14.217	25.334	2351626
Total Precip (Inches)	0	1.009	2.550	2.967	4.349	13.819	2351626
IQR Precip (Inches)	0	0.004724	0.034655	0.158315	0.238235	2.434279	2351626
Quantity (kGal)	0.1	2.6	4.3	6.537	7.4	1275.8	2351626
Payment ($)	8.426	19.118	27.436	57.746	53.074	17577.749	2351626

Note: Heavy Water Appliances include pool, hot tub, sprinkler system, fountain, and car wash station/area. In the real estate industry, a full bathroom requires a sink, a tub, a shower, and a toilet. If it only has 3/4 of fixtures, it will constitute a 3/4 bathroom, and only 2/4 fixtures will constitute a half bathroom.

MLE Estimation Result

Given the model with the interaction terms, the MLE estimates the parameters (β_1', β_2', β_3', β_4', β_5', c, c_α, c_ρ, σ_η, σ_v) from data ($X_i, Z_i, X_{\alpha,i}, Z_{\alpha,i}, X_{\rho,i}, p_k, q_k, A_k, I_i, w_i$), with a total of 18 parameters. All data measuring price or payment are scaled from nominal value to real dollar value in January 2017 using the Federal Reserve Economic Database Gross Domestic Product (GDP) deflator from St Louis Federal Reserve Bank. [38] The estimation results are listed in Table A5:

Table A5—: MLE Estimation Results

Parameter Name	Estimate	Standard Error	Parameter Name	Estimate	Standard Error
Bathroom	1.16	(9.20E-04)	Average High Temp	0.00218	(2.04E-05)
NDVI	1.19	(0.011)	IQR High Temp	-0.0189	(1.46E-04)
Constant	0.731	(0.0017)	Total Precipitation	-0.941	(1.34E-04)
Price * bedroom	0.0486	(3.94E-04)	IQR Precipitation	0.12	(0.005)
Price * NDVI	-0.0404	(0.00465)	Income * Heavy Water Appliances	-0.0718	(4.81E-04)
Price * Avg High Temp	-0.0168	(2.47E-05)	Income * Bedroom	-0.0158	(5.34E-05)
Price * Total Prcp	-0.0363	(4.23E-04)	Income * NDVI	-0.0692	(3.91E-04)
Price	0.688	(0.00193)	Income	0.162	(1.67E-04)
σ_η	2.56	(8.45E-04)	σ_v	4.71E-04	(2.25E-03)

Note: Price * bedroom represents the interaction term inside α, measuring how the price effect is changed through the number of bedrooms. The same goes for other interaction terms for price and income.

Most results of MLE fit the intuition qualitatively. The standard error is calculated by the inverse of the matrix of the sum of the outer products of the observation-by-observation gradient of the log-likelihood for each household evaluated at the maximum likelihood parameter estimates, a method introduced by Hall, Hall and Hausman (1974).

Price Elasticities of Income Strata

From Section III.C, I have shown the counter-intuitive result that high-income households are more elastic when the weather is dry. For high elasticity households to pose a "threat" to the pricing design, they will need to be high quantity users as well. I then plot the price elasticity on quantity for all strata.

From Figure A5, there are plenty of higher-income, high elasticity households with large quantity. In the highest income stratum, 14.4% of the households have the elasticity lower than the overall median and have the quantity to be larger than 20k gallons. In addition, there are a decent number of lower-income households that have pretty high consumption levels with not very high elasticities. Both the mismatches between income and quantity, and income and elasticity, explain the reduced-form evidence from 2, and weather is one of the factors that causes these mismatches. This can be a potential mechanism explaining the findings from previous literature [39] where they conclude that IBP

[38] See Fred-GDPDEF

[39] Ruijs, Zimmermann and van den Berg (2008), Echeverri (2023), and Wichman (2024)

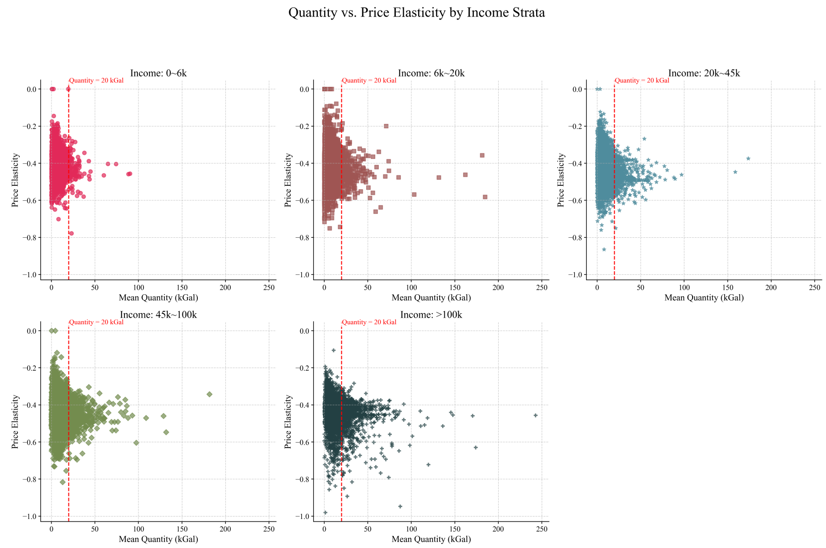

Figure A5. : Price Elasticities v. Mean Quantity by Income Strata

optimization will likely favor higher income strata. Essentially, I would like the utilities to be aware of the price sensitivities of higher income strata, as the higher paying consumers are not actually "reliably" contributing towards the revenue, as well as the existence of the low elasticity-high quantity households in the lower income strata, as these consumers who should be benefiting from the equity goal, could potentially suffer from other competing policy goals.

One important caveat for this study (and further discussion of distributional results) is that the household income I calculated actually stemmed from the correlated house value. Even though it has been proven to be generally true[40], it is a strong assumption to assume the same distribution of house value and household income within the same zipcode. However, later during the counterfactual analysis, I conduct a specification with only the newly built houses, such that their appraisal value will be generally higher, and still find that welfare for the lowest income stratum is much lower compared to other strata. The result can be found in Section A.A4.

A3. Ramsey Pricing Model

STATUS QUO REVENUE AS THE THRESHOLD

I will explore the validity of using the status quo revenue as the lower bound for counterfactual revenue in this section. Due to the lack of detailed cost breakdown data, it would be difficult to establish a supply-side cost model. However, by using aggregate cost information, I can make some inferences. In order to study the financial viability

[40]See Zhang (2016), and Kim (2020)

of Austin Water, I need to ensure that total revenue can cover total costs. However, total water service revenue is only a portion of Austin Water's overall revenue stream. Furthermore, the revenue I observe in this data represents only single-home residential water revenue, which itself is only a portion of total water service revenue. Nonetheless, water service revenue is typically the largest revenue stream (> 70%), and single-home housing usage usually constitutes the largest part of total water service revenue (> 50%). In particular, single-home housing usage is most susceptible to precipitation variation because multi-home housing is typically in condominiums, which do not have lawns, and commercial usage usually does not vary with changes in precipitation. Therefore, it is valid to use single-home housing to study the impact of weather, and compare the total revenue from single-home housing usage to some threshold to measure the revenue risk faced by the utility due to weather variation.

However, the lack of detailed cost information makes it hard to estimate the cost generated solely by single-home housing. The total costs from water services include operation, labor, debt service requirements, and funds transferred to the city government. None of these can be separated and attributed solely to single-home housing usage.

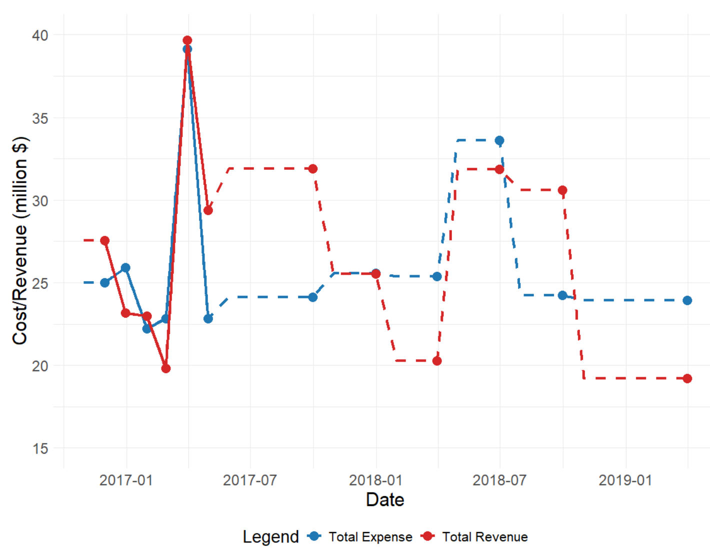

Figure A6. : Total Cost and Revenue from Water Service

Note: Austin Water releases its financial report irregularly, but on average, every 3 months. During some time between January 2017 and May 2017, they released the report every month. Hence, some data have a monthly frequency, and some do not.

Looking at the seasonal trends from quarterly financial reports of water service revenue in Figure A6 and cost, I can see that the cost for the whole year is relatively stable. The only peak is usually around the summer months when the water supply is lower and requires more water from the reservoir. The main reason for negative profit in water service is that revenues in some months fall below the typical cost level during non-peak months. Therefore, the financial risk faced by the utility is largely due to revenue volatility between months. Therefore, it is reasonable to set up a revenue lower bound for comparison with counterfactual revenues. I could set up an intricate cost model to estimate this lower bound, but the status quo revenue (which is the revenue of the year 2019, and the weather variation in 2019 was quite standard) serves as a good enough benchmark.

Simulate Weather Distribution

It is tricky to evaluate E_Z without the knowledge of the real distribution of Z. This is due to 1) if limited to recent years of Z data, there are too few data points to empirically generate the real distribution of Z, and 2) if expanding to long-term data of Z, the weather data 30 years ago does not share the macro climate trends of the recent weather data. Unfortunately, truly predicting the weather is out of the scope of this paper, but I offer an alternative solution to estimate the optimal price based on the prior of the utility. In addition, in one of the specifications, I used a Monte Carlo method to simulate weather prediction errors and the optimal result does not deviate that much.

The counterfactual weather is generated from a prior shift from the utility. If the utility thinks next year's weather is on average more rainy, then $\zeta_1 > 0$ and vice versa. If the utility thinks next year's weather is on average higher in variance, then $\zeta_2 > 1$. The utility will treat this counterfactual Z as the static weather and then calculate the optimal price.

Adding a Monte Carlo small log-normal disturbance will simulate the prediction error and the result is more realistic. I generate a small log-normal disturbance with $\mu = 0$ and $\sigma = 1$. The $\sigma = 1$ is calculated from the standard deviation of the strictly positive precipitation data from the recent 5 years (2014-2018). This means the utility first predicts a general trend of precipitation data, it could be either $Z' = Z \pm \zeta_1$ or $Z'/Z = \zeta_2$, and then in order to account for prediction errors, Z' is perturbed by a small log-normal error term to simulate the stochasticity of Z' based on the utility's prediction.

The Monte Carlo process is calculated as follows: for each simulation s, the utility will calculate both the objective $(CS(Z^s) - \lambda \cdot \max(0, C - R^s(Z^s))$ and the conservation constraint $(P(\sum q(Z^s) \leq \bar{Q}) \geq 0.95)$ and take the average of all numbers of simulations. Essentially, the utility is maximizing over the average objective for all the simulations, while making sure the average quantity for all simulations is below the threshold. One might question why not average over all simulated months, with a total number of simulated data points to be $s \cdot 12$. It is hard to imagine the utility optimizing on a monthly basis, even though it could generate more data points. On the other hand, it is reasonable to assume that the utility would need to check the projected overall welfare and projected revenue condition by the end of their financial year. Averaging the yearly objective for all simulations essentially calculated the expected objective for the utility under the weather stochasticity. Due to the long running time, I eventually chose $s = 25$ as a start. The resulting optimal prices do not change significantly compared to Figure 9 and Figure 10. This showcases that the baseline model gives a consistent measurement of the first moment of weather stochasticity.

A4. Counterfactual Analysis

COUNTERFACTUAL NDVI

As NDVI measures the health of vegetation, it would not make sense to keep it unchanged through shifting counterfactual precipitation. To isolate the effect of precipitation on NDVI, I conducted an OLS using the following formula, where P represents precipitation and T represents daily max temperature:

$$NDVI = \beta_0 + \beta_1 P_{\text{sum}} + \beta_2 P_{\text{iqr}} + \beta_3 T_{\text{mean}} + \beta_4 T_{\text{iqr}} + \beta_5 I$$

Note that both the mean and IQR here refer to the within-month mean and IQR. For this paper, I focus on the variance of the monthly sum of precipitation (P_{sum}); therefore, while keeping all the other variables unchanged, the parameter of interest will be β_1. The OLS results are as follows:

Table A6—: OLS Estimation Results of Precipitation on NDVI

Variable	Estimate	Std. Error	t value	$\Pr(> \|t\|)$
(Intercept)	0.1149	0.0010	92.940	$< 2\text{e-}16$ ***
P_{sum}	0.0080	0.000045	177.803	$< 2\text{e-}16$ ***
P_{iqr}	-0.0118	0.0004	-29.460	$< 2\text{e-}16$ ***
T_{mean}	0.0030	0.000012	256.693	$< 2\text{e-}16$ ***
T_{iqr}	0.0013	0.000032	40.768	$< 2\text{e-}16$ ***
I (in millions)	1.449	1.3	114.940	$< 2\text{e-}16$ ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Multiple R-squared: 0.127, Adjusted R-squared: 0.127
F-statistic: 3.935e+04 on 5 and 1,356,918 DF, p-value: $< 2\text{e-}16$

This regression captures the impact of precipitation on vegetation health ecologically. The only control variable included is income to capture the size of the lawn for each household. This means for every change in precipitation in inches, the NDVI will increase by 0.008, keeping demand and income the same. Through the calculation of counterfactual NDVI, I can calculate the counterfactual demand, welfare, etc.

OPTIMAL PRICE PROCEDURE

Given the counterfactual weather and NDVI, household characteristics, and income, I will use Equation 12 to solve for the optimal price of 14 parameters. All parameters are part of IBP; therefore, I set all prices to be larger than 0, and force each step to be increasing for both marginal prices and fixed payment (meaning the difference between each tier will at least be $0.01). The optimization algorithm I used is COBYQA, a derivative-free optimization solver designed to supersede COBYLA to solve for the bounded optimization process. In essence, COBYQA is a trust-region SQP method based on quadratic models obtained by underdetermined interpolation.[41]. Of all the constrained optimization algorithms, COBYQA consistently can provide reliable results as long as the initial value is within the valid range with respect to constraints.

Since I am dealing with price optimization, I do not require very refined results. I chose the initial searching radius of the algorithm to be 1.0, and the final radius for convergence tolerance to be 0.01 (as the lowest price difference will be in cents). I also utilize putting the loss function in the objective function in addition to the conservation constraint to avoid optimizing on 2 constraints at the same time (since otherwise, without the loss function, both constraints will need to be implemented). I choose to set the initial value as if each price jump is the same between tiers to intensify the certain benefit of more salient price jumps between certain tiers. Since COBYQA is good at finding relatively local results, I choose the initial value to be close to the status quo IBP.

Aside from weather stochasticity, I set up a Monte-Carlo simulation for ε using the result from Table A5. Note that there are two levels of stochasticity with unobserved taste (η) and perception error (ν). The perception error is purely random and not controlled by the household, so for each household for each month, I will generate a simulation for ν_{mh}. As for η, I assume each household has an unchanged $\bar{\eta}_h$ across all months to avoid overfitting and simulate the results. I have tried using η_{mh}, the counterfactual welfare generates very little difference between specifications.

Price Result - Mean Shift

Since the 5-tier IBP includes both marginal prices and fixed payments, I argue the most effective way to visualize the pricing structure is by plotting total payment (in dollars) versus quantity (in thousand gallons). Note that all plots are truncated at 50 k gallons to focus on the changes around the majority of the data points. In practice, a decent number of consumers use significantly more than 50 k gallons. In this subsubsection, I will only focus on ζ_1 and only showcase the result of $\zeta_1 = -0.25$ and $\zeta_1 = 0.25$. The results of non-extreme weather conditions in between fit the general trend and are omitted for clarity. All price plots include three benchmarks derived from the status quo. The status quo price (shown with a dot-dash line) reflects the current pricing structure used by Austin Water. The status quo flat price is constructed by averaging the five marginal prices and five fixed payments, resulting in a linear pricing structure. I also include the empirical cdf of the quantity to showcase the change in distribution.

From Figure A7, when $\zeta = -0.25$, weather pushes the demand curve to the right, but the binding conservation constraint actively curb high quantity users, making the price for the higher tier much higher, and quantity distribution for higher quantity (> 20k Gallons) shifts to the left. The rising price for the higher tier causes welfare loss for the high quantity users, and the existence of the high-quantity-low-income users generating the regressive welfare effect from pricing.

[41]The algorithm is developed by Ragonneau (2022), Ragonneau and Zhang (2025)

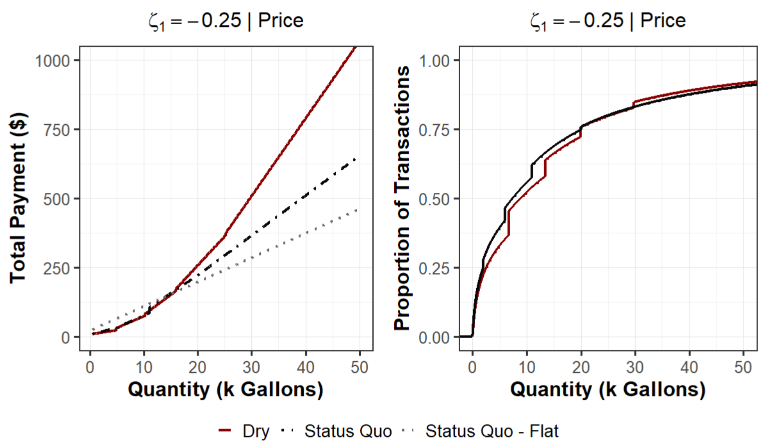

Figure A7. : Optimal Prices - $\zeta_1 = -0.25$

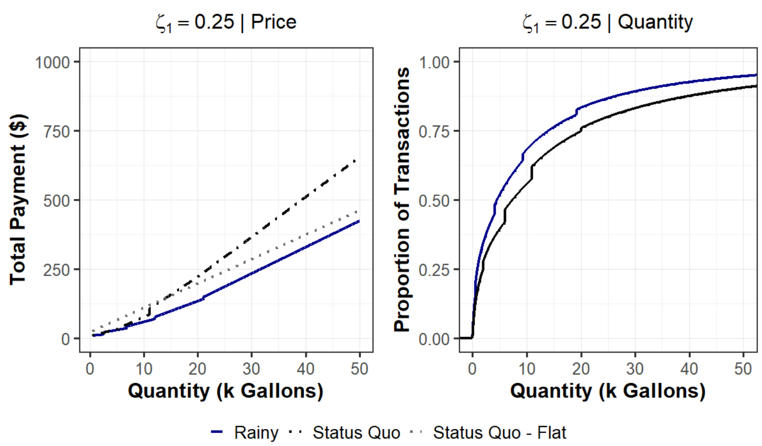

Figure A8. : Optimal Prices - $\zeta_1 = 0.25$

From Figure A8, when $\zeta = 0.25$, weather pushes the demand curve to the left. Since the revenue loss is less in priority for the utility ($\lambda = 0.5$), and the existence of high-quantity-low-income users push the price downward for all tiers. Even with the price decrease, the demand curve still shifts to the left, as evidenced by the empirical cdf.

PRICE RESULT - VARIANCE SHIFT

When changing ζ_2 (the ratio of weather standard deviation compared to the status quo), I will only focus on the extreme cases of $\zeta_2 = 0.75$ and $\zeta_2 = 1.25$.

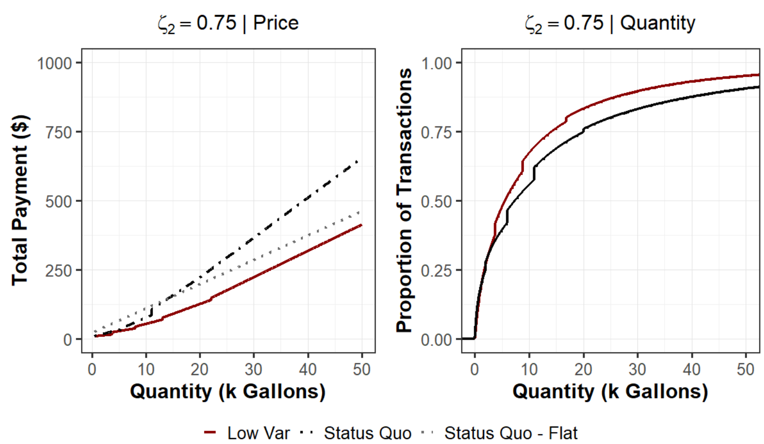

Figure A9. : Optimal Prices - ζ_2

From Figure A9, when $\zeta = 0.75$, the quantity distribution shifts towards the median. However, since the status quo (and the counterfactual) quantity distribution is right-skewed, the quantity in general decreases. This will result in a similar situation of $\zeta_1 > 0$ where the revenue losses are more pressing, which will in turn result in an increase in prices across all tiers.

From Figure A10, when $\zeta = 1.25$, the quantity distribution becomes more spread out. Due to the right-skewness of the data, the quantity in general increases. This will result in a similar situation of $\zeta_1 < 0$ where the conservation constraint is binding. However, what is different from the cases of $\zeta_1 < 0$ is the decreasing quantity in even drier months, making the price for lower tiers higher due to the financial risks during the now even drier months.

WELFARE RESULTS OF DIFFERENT BUILT YEARS - ROBUSTNESS CHECK

One of the concerns regarding the welfare result is that, since I am essentially equating house value to household income. There could be some mis-categorization where a household built early with a low appraisal value, but actually has a higher income. I

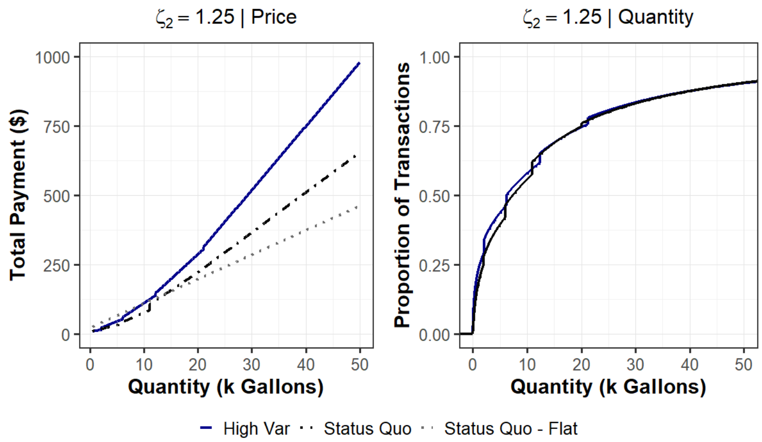

Figure A10. : Optimal Prices - ζ_2

am going to test the counterfactual analysis with the welfare from the optimal price in shifting mean and extend the result from Figure 9 with different segments of the households cohort. I will test two specifications: 1) limiting to only the houses built after 2000, which is the top 27.4% of the entire cohort, and 2) limiting to only the houses built after 1982, which is the top 50% of the entire cohort. Since the lowest income stratum will have the largest welfare swing, I will only limit the welfare result of that stratum.

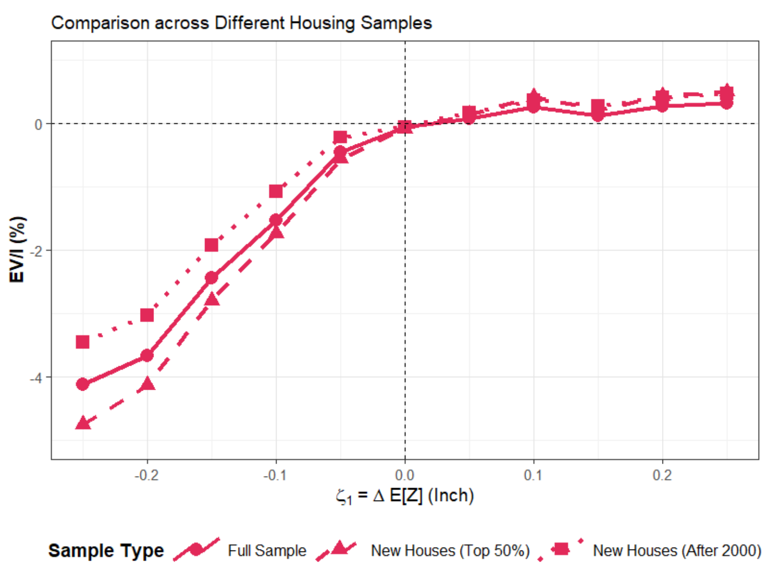

Figure A11. : Welfare from Status Quo Price in Shifting Mean ($\zeta_1 \in [-0.25, 0.25]$)

When comparing to the full sample, both specifications still generate comparable welfare results, and limiting to houses with the built year in the top 50th percentile even generates lower welfare. When $\zeta_1 = -0.25$ (the demand shifts to the right), the full sample has an average EV/I for the lowest income stratum of -4.11%, while 1) limiting to only the houses built after 2000 (the 27.4 th percentile) has an average EV/I for the lowest income stratum of -3.45%, and 2) limiting to only the houses built after 1982 (the 50 th percentile) has an average EV/I for the lowest income stratum of -4.75%. The fact that limiting the sample to above the median generates even lower welfare, showcasing that the way of estimating income (namely, extrapolating the zip-code level income by house value distribution) is largely accurate, and the welfare result from Figure 9 is robust.

As for the shadow cost of the policy constraints, when $\zeta_1 = -0.25$, the shadow cost of the conservation constraint incurs an average welfare loss of \$60.92 per household per month for the lowest stratum for the full sample[42]. For specification 1), which is limited to houses built after 2000, the same weather conditions and the welfare result imply an average welfare loss of \$60.89 per household per month for the lowest stratum. For specification 2), which is limited to houses built after 1982, the same weather conditions and the welfare result imply an average welfare loss of \$72.94 per household per month for the lowest stratum.

When $\zeta_1 = 0.25$, the shadow cost of the revenue constraint incurs an average welfare loss of \$63.65 per household per month for the lowest stratum for the full sample. For specification 1), the same weather conditions and the welfare result imply an average welfare loss of \$63.08 per household per month for the lowest stratum. For specification 2), the same weather conditions and the welfare result imply an average welfare loss of \$70.82 per household per month for the lowest stratum. The fact that the welfare damages from the shadow cost of the policy constraints between specification 1 and the full sample are so similar showcases the robustness of my results, and using the house appraisal value to back out income won't bias the results.

I calculate the shadow costs of the policy constraints for the two specifications as the robustness check of the result of shifting variance. When $\zeta_2 = 0.75$, the shadow cost of the conservation constraint incurs an average welfare loss of \$67.19 per household per month for the lowest stratum for the full sample. For specification 1), the same weather conditions and the welfare result imply an average welfare loss of \$65.72 per household per month for the lowest stratum. For specification 2), the same weather conditions and the welfare result imply an average welfare loss of \$74.30 per household per month for the lowest stratum. When $\zeta_2 = 1.25$, the shadow cost of the conservation constraint incurs an average welfare loss of \$74.87 per household per month for the lowest stratum for the full sample. For specification 1), the same weather conditions and the welfare result imply an average welfare loss of \$75.96 per household per month for the lowest stratum. For specification 2), the same weather conditions and the welfare result imply an average welfare loss of \$84.82 per household per month for the lowest stratum.

[42]See SectionV.C

STATUS QUO PRICE - WEATHER CHANGE

In order to separate the welfare loss from weather itself and the welfare loss from the shadow cost of the policy constraints, I calculate the welfare using the status quo price vector $\mathbf{p}_0$ under all counterfactual weather conditions ($\zeta_1 \in [-0.25, 0.25], \zeta_2 \in [0.75, 1.25]$). Here are the welfare result for the mean shift ($\zeta_1 \in [-0.25, 0.25]$) and variance shift ($\zeta_2 \in [0.75, 1.25]$):

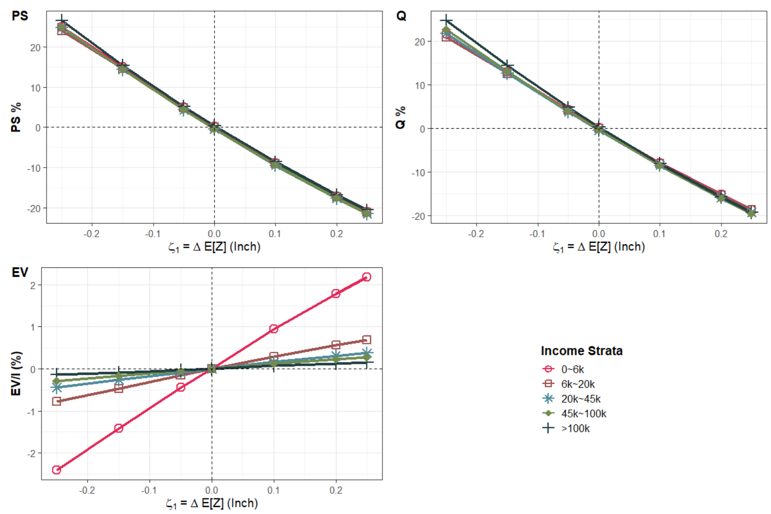

Figure A12. : Welfare from Status Quo Price in Shifting Mean ($\zeta_1 \in [-0.25, 0.25]$)

I can see the clear shift of the demand curve through the weather and how it's affecting both total quantity and total payments from all strata for both cases of shifting mean and variance. Without the optimal pricing procedure considering the policy constraints, it will be difficult to curb consumption when $\zeta_1 < 0$ and lower bound the revenue when $\zeta_1 > 0$. However, the upside of not including the policy constraint is an increase in welfare, in particular, a significant increase in welfare for the lowest income stratum.

HIGH QUANTITY CONSUMERS

There are plenty of high quantity consumers, but what truly makes the equity goals of the optimal price hard to achieve is the existence of high quantity consumers in lower tiers. I have already explored the existence of these consumers in the data, but here I would like to specifically show the welfare loss incurred by these customers in the lowest stratum (0$\sim$6k) and compare it to the highest stratum (>100k). I will focus on the specific weather pattern of $\zeta_1 = -0.25$ where the welfare for the lowest stratum is the lowest.

From this picture, I can clearly see that the magnitude of welfare losses is higher in

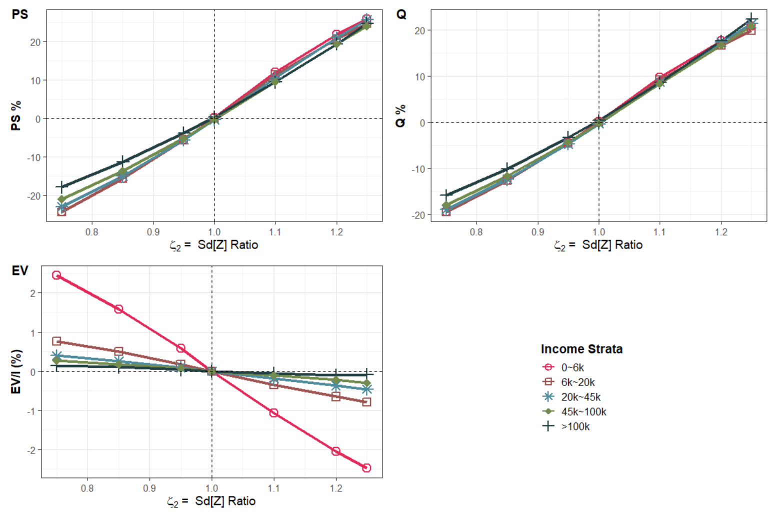

Figure A13. : Welfare from Status Quo Price in Shifting Variance ($\zeta_2 \in [0.75, 1.25]$)

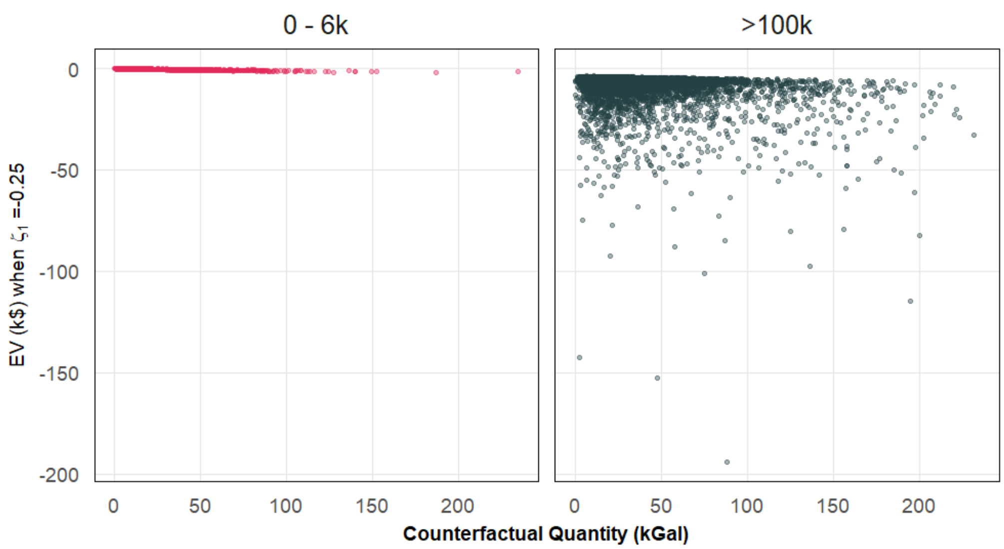

Figure A14. : EV when - $\zeta_1 = -0.25$, $0 \sim 6$k and >100k strata the higher income strata across different quantity levels. However, if I measure EV/I, which is a fairer measurement of welfare across income strata:

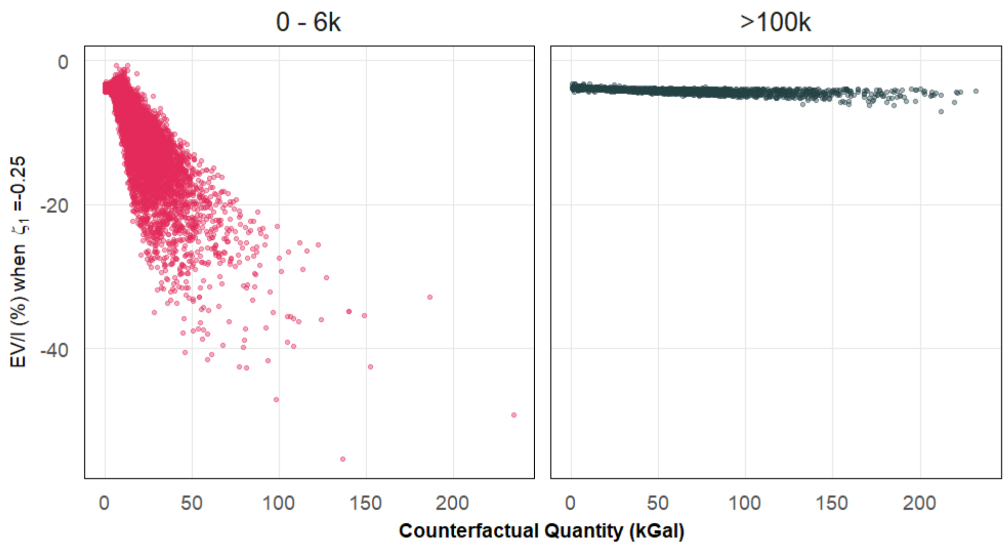

Figure A15. : EV/I when - $\zeta_1 = -0.25$, $0 \sim 6$k and >100k strata

Even though the higher income stratum loses more welfare, the high quantity users (those higher than 20k gallons), and their relatively low elasticities, cause lower welfare in extreme weather conditions for the lowest income stratum.

To separate the heterogeneities of the quantity, and since income turns out to be a problematic predictor of quantity. I performed a reduced-form logit regression to figure out the source of the heterogeneity. I define high quantity transactions as > 20k gallon and only limit to the lowest income stratum. I include some predictors from the demand model, as well as variables like HVAC category (a categorical variable labeling the size of the house), lawn percentage (a continuous variable measuring the percentage of the lot size not counted as living areas), and the interactions with precipitation. Here is the equation and the result.

$$\begin{aligned}\mathbb{1}\{\text{High Quantity}\} = {} & \text{house value} + \text{precipitation} + \text{NDVI} + \text{bathroom} + \text{bedroom} \\ & + \text{spa area} + \text{heavy water app} + \text{HVAC category} + \text{lawn percentage} \\ & + \text{precipitation} \times \text{house value} + \text{precipitation} \times \text{NDVI} \\ & + \text{precipitation} \times \text{bathroom} + \text{precipitation} \times \text{bedroom} \\ & + \text{precipitation} \times \text{spa area} + \text{precipitation} \times \text{heavy water app} \\ & + \text{precipitation} \times \text{HVAC category} + \text{precipitation} \times \text{lawn percentage}\end{aligned}$$

Table A7—: Logit Regression Model of Predicting High Quantity for Low-Income Households

Main Effects				Interaction Terms			
Predictor	**Est.**	**Std. Err.**	**z val.**	**Predictor**	**Est.**	**Std. Err.**	**z val.**
(Intercept)	-1.031***	0.138	-7.458	Precipitation : house value	0.002***	0.000	21.639
house value (k$)	-0.004***	0.000	-38.242	Precipitation : NDVI	-0.369**	0.125	-2.965
Precipitation (Inch)	-2.345***	0.138	-17.061	Precipitation : bathroom	-0.066***	0.017	-3.902
NDVI	5.294***	0.129	41.048	Precipitation : bedroom	0.002	0.015	0.112
bathroom	0.209***	0.019	10.892	Precipitation : spa area	-0.001***	0.000	-10.222
bedroom	-0.016	0.016	-0.996	Precipitation : heavy water app	0.110	0.152	0.723
spa area (sqft)	0.003***	0.000	13.203	Precipitation : HVAC Med-Low	-0.104**	0.033	-3.101
heavy water app (sqft)	-0.091	0.199	-0.458	Precipitation : HVAC Med-High	0.037	0.039	0.967
HVAC Med-Low	0.414***	0.032	12.944	Precipitation : HVAC High	0.128**	0.048	2.691
HVAC Med-High	0.696***	0.042	16.721	Precipitation : lawn percentage	0.472**	0.155	3.036
HVAC High	1.009***	0.057	17.739				
lawn percentage	0.212	0.156	1.361				

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1. Est. = Estimate, Std. Err. = Std. Error, z val. = z value.

For households in the lowest-income stratum. Higher living area, more attention to lawn (NDVI), more bathrooms, and larger pool/hot tub areas (in sqft) are all strong predictors of increased water use. House value has a negative coefficient. Based on my assumption to correlate house value with house income, and the conclusion that income is not a strong predictor of heterogeneity, this is not entirely surprising. This showcases that some households with low house value (which means low income) have very high quantities. The influence of these factors changes with precipitation: the impact of vegetation is most pronounced in dry conditions, while the effect of lawn percentage and the consumption gap between the largest and smallest homes becomes more significant in wetter climates. Therefore, for the cases of dry weather conditions, where welfare loss is the largest, NDVI is a good predictor for the heterogeneities of high water consumption, and zeroscaping is a natural additional policy to remedy the impact from NDVI. Further studies could be extended towards the impact of bathrooms and larger pool/hot tub areas and their related policy impacts.